Responses to Recurring Issues Related to North Dakota's Computer Modeling of Sulfur Dioxide in CAA PSD Class I Areas

final

August 3, 2007

North Dakota Department of Health
Environmental Health Section
Air Quality Division
918 E. Divide Ave.
Bismarck, ND 58501-1947

Ambient Air Quality Monitoring Site Located in the North Unit of TRNP

Responses to Recurring Issues Related to North Dakota's Computer Modeling of Sulfur Dioxide in CAA PSD Class I Areas

final

August 3, 2007

Prepared by Martin R. Schock
North Dakota Department of Health
Environmental Health Section
Gold Seal Center, 918 E. Divide Ave.
Bismarck, ND 58501-1947

Purpose

The State of North Dakota (State) and the North Dakota Department of Health (NDDH) have developed and implemented an air quality modeling protocol using the U.S. Environmental Protection Agency (EPA) approved Calmet and Calpuff models. The protocol evolved in a process between the State and EPA to resolve different inputs for these models and interpretation of outputs by these models.

The EPA orally agreed to a NDDH air quality computer modeling protocol on April 28, 2004. EPA later expressed a preliminary view that the NDDH's modeling techniques are consistent with the federal Clean Air Act by letters dated June 30, 2005 and August 17, 2006.

Since April 2004 and June 2005, some issues pertaining to the NDDH's protocol have recurred, in spite of verbal consent by EPA to proceed with modeling and preliminary written approval. This paper was prompted by a computer PowerPoint presentation/report prepared by Patrick. L. Hanrahan, Portland, Oregon, which was released August 16, 2006. Hanrahan's report [1] was prepared for the Dakota Resource Council.

Hanrahan's PowerPoint echoes elements of a computer email memorandum dated April 21, 2004, over names of EPA regional modelers, [2] excluding Region 8, Denver, Colorado, and PowerPoint presentations/reports prepared by Region 8's Kevin Golden. Golden's first report [3] was presented at an EPA Region 4 workshop of regulatory modelers held in March 2005; his second report [4] was presented at a national workshop of modelers held in May 2005.

All PowerPoint reports relate (1) to computer modeling protocols by the NDDH and by EPA Region 8 and (2) to results of the implementation of those protocols.

These reports did not provide any new information not otherwise already assembled in a NDDH hearings docket. These reports and the email memorandum fail to justify or defend modeling methods as used by EPA Region 8 in its 2002 and 2003 draft modeling in context of the federal Clean Air Act, EPA's implementing rules and regulations, court decisions or scientific principles.

---

[1] See http://www.drcinfo.com/

[2] Several issues in the email memorandum first appeared on pages 8 – 26 in Exhibit 57. This exhibit was prepared by Region 8, is dated May 24, 2002, and is titled *EPA Comments on NDDOH's Proposed Determination Regarding the Adequacy of the SIP to Protect PSD Increments for Sulfur Dioxide*.

[3] See http://www.epa.gov/region4/air/modeling/workshop.htm Page 10 lists "Issues of Concern". See table S1.

[4] See http://cleanairinfo.com/modelingworkshop/presentations/rls_dispersion.htm

Table of Contents

Acknowledgments

Assistance in preparation of this paper was provided by the following NDDH staff.

Dan Harman	Air Quality Data (Parts 2 and 5)
Tom Bachman	Emissions and Coal Data (Parts 2 and 6)
Steve Weber	Maps (Part 2) and Computer Modeling (Part 5)

Preface

P.1 The State and the NDDH have a long history of experience with PSD.

The State and NDDH have extensive experience in the federal Clean Air Act (CAA) PSD program, including air quality modeling. [5]

> "North Dakota, through the North Dakota Department of Health [State or NDDH], has an EPA-approved SIP, see 40 C.F.R. Part 52, §§ 1820-1835, under the Clean Air Act, and has administered the PSD provisions of the Act since the PSD law was implemented shortly after its passage by Congress in 1977."
>
> As a result of "the mid-East oil crisis of the 1970s, North Dakota experienced an unprecedented energy development boom, when its oil, natural gas, and lignite coal resources were explored, permitted and developed as never before (or since)."
>
> "The permitting of a large number of these coal-fired facilities during this period placed North Dakota at the forefront of implementing and applying newly enacted and promulgated PSD statutes and regulations in a NAAQS attainment area." [6]
>
> "Coal-fired [and natural gas processing] facilities permitted and constructed shortly after the minor source baseline date of December 19, 1977, were determined to have consumed all available sulfur dioxide increment in the state's PSD class I areas based upon computer air quality modeling methods used at that time."
>
> "Between 1982 and 1993, numerous new facilities were permitted in North Dakota [by the NDDH] for construction based on federal land managers (FLM) determinations under 42 U.S.C. § 7475(d) that they would have no adverse impacts on air quality related values (AQRVs). See, e.g., 47 Fed. Reg. 30,222 (1982), and Fed. Reg. 13,639 (1993). Two state facilities continue to operate under these FLM certifications of no adverse impact."

After 1993 and until 1999, no additional PSD new source review (NSR) mesoscale modeling for air quality impacts on PSD class I areas was done by the NDDH. A NSR occurs prior to source construction and prior to a history of source actual occurring emissions; so compliance with NAAQS and PSD increments under NSR is usually determined through air quality modeling.

---

[5] See Exhibit 155 titled *August 29, 2005 Findings and Conclusions and September 7, 2005 Determination*, pages 3, 4 and 5.

[6] See Exhibit 126 titled *A case history of the North Dakota PSD program*. Environmental Science & Technology, Vol. 16 (1982). No. 7, by Myron F. Uman, National Research Council.

P.2 The NDDH has completed a periodic review of the adequacy of its SIP.

During the years from 2001 to 2004, the NDDH and EPA Region 8 twice conducted separate computer modeling of sulfur dioxide emissions to assess the on-the-ground status of attainment of CAA PSD class I 24-hour and 3-hour increments. During these years, the NDDH conducted a periodic review hearing under North Dakota law and 40 CFR § 51.166(a)(4) pertaining to the legal and technical framing for requirements and for discretionary choices in its modeling practices.

The State Health Officer's signed findings after the 2002 and 2003 periodic review hearing are dated September 8, 2003. These findings directed the NDDH to continue to engage EPA to resolve outstanding issues through written agreement and negotiations. [7]

On February 24, 2004, ***the State of North Dakota and EPA entered into a Memorandum of Understanding (MOU) that created a process to resolve differences in their respective modeling practices***, [8] and the State agreed to conduct additional modeling. This MOU also described several avenues of State discretion under the federal CAA when conducting air quality modeling. ***EPA orally approved the State's MOU modeling protocol on April 28, 2004***; copies were forwarded to EPA during May 2004 and November 2004. The NDDH subsequently held another periodic review hearing in April 2005 regarding its MOU modeling protocol and other MOU-related documents comprising a ***draft version*** of North Dakota's SO2 PSD Air Quality Modeling Report.

In a letter dated June 30, 2005, EPA states "***[O]ur preliminary view is that the modeling techniques that you used in your PSD analysis are consistent with the Clean Air Act.***" (emphasis and italics added) Those modeling techniques and practices are described in the State's MOU Protocol, which is Addendum B attached to a ***final version*** of North Dakota's SO2 PSD Air Quality Modeling Report. In another letter dated August 17, 2006, EPA states "[I]n June 2005 ... we were inclined to agree that North Dakota currently has the discretion to interpret its State Implementation Plan to allow such [modeling] practices." At the time this report was completed, EPA's final approval of the State's protocol and final report is pending.

This paper is a condensed summary of selected aspects and details included in the State's periodic review. Information presented in this paper is taken from exhibits and transcripts included in the NDDH's hearings docket, except for additional literature references, additional sulfur dioxide emissions data and sulfur dioxide actual ambient concentrations, additional model results uncertainty and accuracy analyses, and an actual-ambient-concentration-based increment consumption analysis.

---

[7] See Exhibit 158 in NDDH's hearings docket titled *North Dakota's SO2 PSD Air Quality Modeling Report*, section 1.0. See also Addendum E to Exhibit 158, page 3, which is the State Health Officer's signed Findings and Conclusions for the NDDH's May 2002 hearing.

[8] See Addendum A to Exhibit 158, section III.

Summary

The State, represented by the NDDH, and U.S. EPA Region 8 used two computer models in tandem to estimate on-the-ground sulfur dioxide concentrations for networks of model receptors in the state's PSD class I areas. Calmet, a diagnostic meteorological model, is used to prepare an hourly three-dimension wind field and other weather data for Calpuff, a pollutant transport, dispersion and depletion model.

The NDDH has not, at any time, used air quality models that were not approved by EPA. The use of the Calmet and Calpuff models in the conduct of modeling (or the execution of modeling protocols) requires a variety of input data necessary for the calculations of model algorithms. The conduct of modeling includes tabulation of model-estimated hourly concentrations in the context of the 3-hour, 24-hour or annual provisions of the CAA for NAAQS and PSD increments.

This paper addresses recurring issues raised (see Purpose) relating to the State's modeling that estimated current (period of concern) sulfur dioxide concentrations and PSD baseline (reference) sulfur dioxide concentrations. The recurring issues focus primarily on model input data for meteorology and source sulfur dioxide emissions and on methods for tabulation of the model-estimated hourly sulfur dioxide concentrations in the context of the PSD increments. The issues also include the roles for actual sulfur dioxide concentrations obtained with on-location monitors and for model performance accuracy testing.

Region 8's May 2005 list of concerns is shown in table S1, page xxi. A directory to sections of this paper that address those concerns is also provided in table S1.

Some of these issues first arose in February 2000 [9] and were carried forward in subsequent meetings and correspondence between the NDDH and EPA Region 8. These and other issues are described or discussed in several exhibits in the NDDH's hearings docket. Most, if not all, of the issues were also discussed by Region 8 in a paper dated May 24, 2002; this paper is Exhibit 57 in the docket. Region 8 provided additional comments in and attached to a letter dated July 1, 2003 (see Exhibit 100).

The State and the NDDH have: a) logged more than five years of labor, b) gathered numerous documents and acquired and assembled large amounts of data, c) researched the history of the CAA and implementing rules, d) researched preambles to rules, court decisions and EPA and FLM decisions, policies, strategies and guidance, e) prepared several comprehensive legal and

[9] See letter in Exhibit 17 from Richard R. Long, EPA Region 8, to Jeff Burgess, NDDH, dated February 1, 2000. This EPA letter provided comments on Exhibit 129 in reply to a letter from Dana K. Mount, NDDH, to Dick (Richard) Long, dated October 21, 1999. Exhibit 129 is a draft report titled *Calpuff Class I Area Analysis for Milton R. Young Generating Station*,

technical documents and f) conducted PSD SIP periodic review hearings where all information was available for review, analyses and comment. This periodic review of a SIP is the first known under the PSD program at 40 CFR § 51.166(a)(4), even though the program was first initiated by EPA in 1974 and then adopted into the CAA by Congress in 1977.

In the midst of this process, ***EPA acknowledged State and NDDH discretion relating to several modeling matters*** in a February 2004 MOU. ***The meteorological, sulfur dioxide emissions and other Calmet and Calpuff input data used by the NDDH for modeling were given oral approval by EPA in April 2004 in advance of implementation.***

In March and May 2005, EPA Region 8 participated in workshops of modeling practitioners (modelers) and listed concerns of the same themes presented in Exhibit 57. However, Region 8 did not bring these concerns to the NDDH in advance of or subsequent to these workshops; nor did it appear, testify or comment in writing during the NDDH's public hearing held April 19, 2005, or the hearing comment period ending June 30, 2005. [10]

This paper summarizes and extends the discussion contained in exhibits and transcripts in the NDDH's docket for public hearings held in 2002, 2003 and 2005. [11] ***The summary that follows reflects State and NDDH findings and experience during the last six years.*** **In sum**:

S1. Preponderant evidence indicates less ambient sulfur dioxide in recent years.

S2. When applying discretion in modeling, the NDDH chose the path established by the CAA, implementing rules, preambles to implementing rules and a court decision, while EPA Region 8 often chose model inputs which exaggerate estimates of worst-case concentrations.

S3. Although data inputs for the Calmet and Calpuff models were improved for the MOU Protocol, uncertainty remains in some input data and model-estimated sulfur dioxide 24-hour concentrations remain inaccurate.

S4. An emphasis on modeling inputs that result in worst-case estimated concentrations is unreasonable, as illustrated by EPA Region 8's absurd modeling results.

S5. Actual concentrations obtained with monitors in the state's PSD class I areas can be used to illustrate annual trends over (or under) an estimated PSD baseline concentration.

---

[10] See Exhibits 141, 150 and 156.

[11] Exhibit 132 is EPA's April 15, 2003, version of its *Guideline on Air Quality Models* (Appendix W to 40 CFR Part 51). EPA has revised and republished Appendix W since that time. This document refers to section and table numbers in Exhibit 132, which may have changed in the current version of the guideline.

S6. From the beginning of the State's periodic review, EPA Region 8, with subsequent support from other federal agency modelers, has promoted consistency and emphasized absurd modeling objectives and methods.

S1. Preponderant evidence indicates less ambient sulfur dioxide in recent years.

Observations:

S1.1 Local oil production in Billings and Dunn County peaked in 1982 and continued on a decline through 2005. State coal production peaked in 1994.

S1.2 Annual total sulfur dioxide emissions:

- ✓ by oil production flares and treaters peaked in 1982, then declined through 1999, and have remained steady thereafter.
- ✓ by electric utilities increased until 1999 and have been decreasing thereafter.
- ✓ by other major sources peaked in 1997 and have been decreasing thereafter.

S1.3 The second highest of annual 24-hour averaged actual ambient sulfur dioxide concentrations:

- ✓ in TRNP-SU peaked in 1989 and declined to lowest values in 2003 and 2004.
- ✓ in TRNP-NU peaked in 1982 and declined to lowest values in 2001 and 2002.
- ✓ at rural Dunn Center peaked in 1983, declined through 1987, and remained steady thereafter.

S1.4 Frequencies (annual number of hours) of actual sulfur dioxide concentrations equal to and larger than 3 ppb (7.9 ug/m3) peaked in 1982 at monitoring sites in the TRNP-NU (Billings County) and at rural Dunn Center (Dunn County).

- ✓ The larger actual sulfur dioxide concentrations during the early and mid 1980s at monitoring sites in the TRNP-NU and at rural Dunn Center were due to sulfur dioxide emitted by nearby oil production flares and treaters.

S2. When applying discretion in modeling, the NDDH chose the path established by the CAA, implementing rules, preambles to implementing rules and a court decision, while EPA Region 8 often chose model inputs which exaggerate estimates of worst-case concentrations.

Observations:

S2.1 Since the early 1980s, EPA's modeling guidelines (e.g., Appendix W to 40 CFR Part 51 and a draft 1990 New Source Review Workshop Manual) have focused on NAAQS or PSD NSR, not PSD periodic review of a SIP. The State and NDDH PSD periodic review is the first known periodic review in the history of PSD since 1974.

S2.2 The PSD increments are permissive amounts of deterioration in air quality over reference concentrations – the baseline concentrations during PSD baseline years – in areas where actual ambient air quality concentrations are less than NAAQS. The increments are changes over baseline concentrations in actual ambient air quality due to changes after the baseline years in pollutant emissions; the increments are not public health standards. Congress intended that increments be administered in ways to prevent adverse impacts on AQRVs, such as soil, water, plants, animals and visibility, in class I areas. Public exposure to cumulative or total air quality is controlled by the primary NAAQS, not by the increments.

S2.3 During a thorough review of the history of the CAA, PSD, EPA's implementing rules and court decisions, as well as preambles to rules and EPA's guidance documents and correspondence, the State and the NDDH identified (a) prescriptive requirements of law and rule (the controlling legal authorities) for modeling and (b) discretionary choices and actions when modeling.

- ✓ Critical provisions of the CAA include monitoring actual air quality and PSD alternate increments. Defining requirements of implementing regulations include the "baseline concentration" and "actual emissions"; EPA's Guideline on Air Quality Models (Appendix W to 40 CFR Part 51) does not address either one. Generally, EPA's use of the phrase actual emissions has contextually meant pollutants actually emitted expressed as rates (e.g., pounds per hour) that are inconsistent with rule-defined "actual emissions".

- ✓ EPA's guidance does not preclude the use of other information, procedures or methods. The guidance is not a substitute for law and regulations, and it does not impose binding, enforceable requirements. Six discretionary options in modeling are described in the February 24, 2004 EPA and State MOU. Four of the options are specific to sulfur dioxide emission inventories and rates. The NDDH implemented these options in its MOU Protocol.

- ✓ Detailed explanations of reasoning in implementation of discretionary options in the MOU are provided in section 3 of North Dakota's SO2 PSD Air Quality Modeling Report and are supplemented with information in this paper.

S2.4 Region 8's 2002 and 2003 modeling methods generally follow the agency's ongoing guidelines (since 1978) for NSR, which emphasize modeling to estimate the extreme, worst-case concentrations. As a consequence, some aspects of Region 8's modeling were inconsistent with the CAA, implementing regulations, preambles to those regulations, a court decision and EPA's guidance as illustrated below.

- ✓ EPA's Guideline on Air Quality Models and its draft (never finalized) 1990 New Source Review Workshop Manual (at pages C.48 and C.49) recommend maximum short-term emission rates for pollutants actually emitted; these documents, including the April 2004 EPA regional modelers' email memorandum, and rates do not reflect or respect the regulatory definition for "actual emissions", which was added to implementing regulations in 1980 (e.g., 40 CFR §§ 51.166(b)(21)(i)&(ii) and 45 FR (August 7, 1980) at pages 52714, 52718 and 52732). In its modeling, Region 8 used 90th percentiles of 24-hour averages of sulfur dioxide hourly CEM emissions for some major sources, and daily averages (total annual emissions divided by 365) for remaining major sources; neither rate is included in the definition, and neither rate is consistent with EPA's guidance.

- ✓ EPA's Guideline on Air Quality Models and other EPA documents recommend model results uncertainty or accuracy analyses. Region 8 did not complete model accuracy tests using all of its chosen model inputs, such as larger current-period emission rates. Its current-period (2000-01) rates are 22.8% larger than the NDDH's current-period rates. And it used 1990 through 1994 meteorology. But actual sulfur dioxide concentrations reflect emitted sulfur dioxide transported by concurrent meteorology. Instead, Region 8 discussed and presented NDDH model accuracy test results as implicitly representing the accuracy of its modeling.

- ✓ Region 8 did not determine a "baseline concentration" for class I areas (or each model receptor in an area) as required by the CAA at § 169(4) and implementing regulations (e.g., 40 CFR § 51.166(b)(13) and also 45 FR (August 7, 1980) at pages 52714, 52718 and 52731). This is because it did not model PSD baseline emissions but instead modeled only increment-affecting sulfur dioxide emissions. A source's increment-affecting emission rate is the difference between its current emission rate and its PSD baseline emission rate.

Recommendation:

When explaining preferred discretionary actions on issues, EPA often cites letters or policy statements, its primary modeling guideline (Appendix W to 40 CFR Part 51) and other documents – all of record since 1974 [12] – in addition to implementing regulations and preambles to these implementing regulations. EPA should not rely on interpretations that are inconsistent with statutes, subsequent rules and court decisions, as well as advanced modeling methods made possible by improvements in technology and data collection.

[12] See, for example, EPA's *SO_2 Guideline Document* dated February 1994 by its OAQPS, EPA-452/R-94-008.

S3. Although data inputs for the Calmet and Calpuff models were improved for the MOU Protocol, uncertainty remains in some input data and model-estimated sulfur dioxide 24-hour concentrations remain inaccurate.

Observations:

S3.1 Calmet and Calpuff algorithms and input data used in modeling connect source sulfur dioxide emissions to estimates of (a) actual ambient sulfur dioxide concentrations and (b) changes in actual ambient sulfur dioxide concentrations between two time lines (e.g., current period and PSD baseline period).

S3.2 Source sulfur dioxide PSD baseline emission rates contain uncertainty. Rates for coal-fired boilers, whether calculated by the State's method or by EPA Region 8's method, are calculated using inexact data for coal-sulfur content, annual coal fired and sulfur dioxide emission factors. Inexact data cause error in model-estimated sulfur dioxide PSD baseline concentrations and calculated deterioration.

S3.3 Available data for oil production flares and treaters during PSD baseline, as well as late-1970s State policies on capturing waste sour natural gas, confound calculation of baseline sulfur dioxide emissions from these sources. ***In the end (section 4.7)***, [13] the inexact baseline data and emission rates for these sources are moot issues because these emissions did not significantly contribute toward modeled consumption of the 24-hour class I increment.

S3.4 Three-point model input and output accuracy assessments were completed.

(a) NOAA/NWS Rapid Update Cycle (RUC) and MM5 data were compared to NWS surface and upper air wind observations within the NDDH's Calmet modeling domain (section 5.2). RUC data appear to agree better than MM5 data with NWS surface wind observations.

(b) Using RUC and MM5 data as input, Calmet output wind fields were compared to independent wind-energy meteorological tower wind data within the NDDH's modeling domain (section 5.2). Whether using RUC data or MM5 as input, Calmet output wind speeds are less than meteorological tower wind speeds for all hours of the day and are also biased toward lower speeds within the range of speeds. But Calmet output wind speeds are less inaccurate when using RUC data.

(c) Calmet–Calpuff-estimated sulfur dioxide concentrations were compared to actual sulfur dioxide concentrations. **In new Calmet–Calpuff performance analyses of estimated sulfur dioxide concentrations (sections 5.3 – 5.8 and Attachment B),** temporal correlation between model-estimated 24-hour

[13] Parenthetical section numbers refer to places in this report where the topic is discussed.

concentrations and actual 24-hour concentrations was examined. Error and bias in the largest 25 model-estimated concentrations were calculated from space(place)-only and space/time pairing with the largest 25 actual concentrations.

- ✗ RUC data do not cause the largest model-estimated concentrations to be less than the largest actual concentrations. And the largest model-estimated concentrations in TRNP when using RUC data are comparable to, or greater than, the largest estimated concentrations when substituting MM5 data for RUC data in the protocol.

- ✗ RUC data provide similar or better correlation between model-estimated concentrations and time-paired actual concentrations when compared to MM5 data. At the rural Hannover monitoring site, the correlation coefficient was higher when using RUC data than when using MM5 data. Correlation was poor at both the NU and SU of TRNP.

- ✗ Error and bias are large. Mean bias, for example, is as large as 5.0 ug/m3, which also is the sulfur dioxide PSD class I 24-hour increment. And mean normalized bias is positive, indicating the models and data inputs are overestimating actual 24-hour concentrations. RUC data provide lower mean normalized bias at the rural Hannover site and more consistent normalized bias among the Hannover and TRNP monitoring sites.

- ✗ Calmet-user choices for mesoscale meteorological data (e.g., RUC or MM5) and for blending these data into NWS surface and upper air observations (i.e., R1 and R2) influence model-estimated concentrations.

S3.5 The new Calmet–Calpuff performance analyses (S3.4(c)), as well as prior NDDH performance analyses, used same-year meteorology, sulfur dioxide emissions and actual concentrations.

- ✓ Emission rates for the new model performance analyses and for the MOU Protocol were rule-defined “actual emissions”, or total annual tons (tons per year) of emitted sulfur dioxide during source operating hours, which reduces to pounds per operating hour. These rates were time constant; i.e., each source’s emission rate was the same for each hour throughout the year.

- ✓ In addition, error and bias did not decrease, or improve, when pairing hourly CEM sulfur dioxide emissions with corresponding hourly Calmet output wind fields as Calpuff input.

- ✓ Given the Calmet-Calpuff performance results, the models and model input data lack skill to adequately match model-estimated current sulfur dioxide 24-hour concentrations with current actual sulfur dioxide 24-hour concentrations. This

is so because the mean normalized bias in the largest model-estimated 24-hour sulfur dioxide concentrations using rule-defined "actual emissions" was not negative, [14] and Calpuff estimated concentrations are, in general, linearly proportional to emissions.

✓ Meteorology, such as wind velocity and mixing height, is four dimensional – x, y, z, and t (time). Locations of stack tops of point sources are three dimensional – x, y and z (stack height); source locations changed after PSD baseline – some were shut down and others started up. Model errors in baseline sources' plume/puff transport and dispersion preceding and during a specific day likely do not match and cancel errors in current sources' plume/puff transport and dispersion. So models also lack adequate skill to protect clean air days (i.e., when using the paired-in-space-and-time method to calculate post-PSD-baseline changes in pollutant concentrations).

✓ The poor model performance is likely due to factors which include: (a) inadequate model algorithms, (b) the stochastic quality of weather in the atmospheric boundary layer, (c) calm or low surface wind speeds and vertical shear (speed and direction) and (d) the lowest reliable calibrated level (LRCL) of sulfur dioxide monitoring instruments.

S3.6 When uncertainty or inaccuracy in model-estimated concentrations are large, it is next to impossible to determine compliance with sulfur dioxide PSD short-term class I increments, such as establishing that the difference between sulfur dioxide 24-hour increment consumption at 4.9 and at 5.1 ug/m3 is real.

✓ Furthermore, Fortran codes for versions of Calmet and Calpuff used by the NDDH in the State's SIP PSD periodic review contained errors. Error fixes caused a significant change in model-estimated sulfur dioxide deterioration for year 2002 in the SU of TRNP.

Recommendations:

Historically, EPA and others have emphasized a discipline for national consistency in modeling protocols (see S4); Region 8's 2002-03 modeling illustrates that this discipline has deterred and impeded expression of modeling uncertainty and accuracy (see S2.4). Modelers should be encouraged to explore and describe the sensitivity and uncertainty of model-predicted (new sources) or model-estimated (existing sources) concentrations in local modeling environments with varied model inputs and with actual concentrations.

---

[14] The NDDH's total 2000-01 sulfur dioxide average emission rates for electric generating units as pounds per operating hour was (a) ~14% less than the total peak rates (as 90th percentiles of daily averages) as pounds per hour (e.g., Exhibit 33, table 3) and (b) exceeded by sums of hourly CEM emissions ~26% of the 17,544 hours during these years (Id., page 3). Note: percentages are given as approximations, since the percentages were not re-computed using the rates in the MOU Protocol.

A wind profiler, located at the rural Dunn Center monitoring site, would obtain hourly horizontal winds from the ground surface through the atmospheric boundary layer. Such data would supplement the NOAA/NWS every 12th hour rawindsonde. These data could be automatically assimilated by NOAA's Rapid Update Cycle mesoscale meteorological model, as are profiler data from other mid-western locations.

S4. An emphasis on modeling inputs that result in worst-case estimated concentrations is unreasonable, as illustrated by EPA Region 8's absurd modeling results.

Observations:

S4.1 Since the late 1970s (e.g., 1979), air quality models and data inputs have been used in NSR to ensure, or essentially guarantee, that a source, once constructed and operating, would not cause or contribute to a violation of a NAAQS or PSD increment. In this approach, uncertainty surrounding modeled concentrations causes users to be cautious, even if conservative inputs mean introducing a systematic bias in modeled results. [15]

S4.2 Consequently, federal agency modelers questing for worst-case model-estimated concentrations has included, for example:

- ✓ Multiple years of meteorology. Three or five years of meteorology are not used to characterize representative pollutant transport and dispersion, but rather to capture and characterize the meteorological transport and dispersion episode(s) causing the worst-case pollutant concentration(s).

- ✓ Maximum, or peak, short-term (24-hour and 3-hour) emission rates for sources constructed after the major source baseline date. Rates larger than 90th percentile rates occur less than 10% of time; maximum, or peak rates, are presumably never exceeded. Furthermore, maximum, or peak, rates are incompatible with the regulatory definition for "actual emissions". EPA Region 8 also used average annual rates for sources constructed before the major source baseline date, which is an inconsistent treatment of sources and expands increment-affecting emissions.

- ✓ Holding the two-year PSD baseline to the two years preceding the PSD minor source baseline date. This practice can distort source PSD increment-consuming emissions when a source's operations during these two years are not representative of normal operations. Alternate years that represent normal operations are allowed by EPA's CAA implementing regulations and preambles to those regulations.

[15] See EPA's *Performance Measures and Standards for Air Quality Simulation Models* dated October 1979 by its OAQPS, EPA-450/4-79-032, page II-21.

✓ In sum, the highest or the second-highest model-estimated concentration due to worst-case emissions during any one year of a multi-year period of meteorological data triggers a possible violation when the concentration is larger than a NAAQS or PSD increment, respectively. The focus of this approach is a single, extreme event and is analogous, e.g., to a 5-year storm event used in flood control management.

An example of these modeling practices is EPA Region 8's draft 2003 report and its estimates of sulfur dioxide PSD class I 24-hour increment consumption. Its modeled sulfur dioxide increment-affecting emissions were 51% of its 2000-01 sulfur dioxide emissions inventory. And Region 8's model estimates of increment consumption for the SU of TRNP ranged between 7.5 and 10.5 ug/m3. Since the Calpuff estimated ambient sulfur dioxide concentrations are, in general, linearly proportional to emissions, Region 8's estimates of actual second-highest (SH) 24-hour concentrations would have ranged between 15 and 21 ug/m3. However, the actual SH 24-hour concentrations for 2000 and 2001 for the SU were 9.4 and 8.8 ug/m3, respectively.

S4.3 The probability of confluence in time of (a) domain-scale meteorology conducive to the largest on-the-ground concentrations, (b) actual peak short-term emissions of each respective source and (c) human occupancy at the place of the largest concentrations is very small.

✓ Worst-case concentrations, generally, relate to acute (one-time peak) exposure by humans when conducting NSR NAAQS analyses rather than to chronic (frequent and less than peak) exposure by AQRVs in PSD class I areas.

✓ From the perspective of the cautious (see S4.1), "[i]t suffices to believe that the existence of a computer prediction of an adverse [or worst-case] situation means that such an outcome is *possible* rather than *correct* in order to take 'action' " [16] without regard as to whether the model predictions are credible (see S2.4 and S4.2).

✓ Although the chances (or probability) – that a new source or new source control strategy will not actually achieve compliance with a NAAQS or a PSD increment – must be acceptable, overly conservative choices for model inputs reflect an attitude favoring minimal chance for underestimating actual

[16] "Testimony of Richard S. Lindzen before the Senate Commerce Committee on 1 May 2001" regarding the climate change debate, page 5.

concentrations. [17] This approach implicitly implements a policy of no tolerance for any residual chance for larger actual concentrations.

S4.4 The emphasis in modeling to capture the improbable, worst-case single event can cause an unreasonable positive bias in model-estimated concentrations. When this is so, abatement strategies may require more source pollutant control than needed, because model performance metric data, such as bias in model-estimated concentrations, have had a minimal and obscure role in Air Quality Management (AQM) decision making.

S4.5 There is no bright line, or threshold, separating unacceptable or inaccurate model-estimated concentrations from acceptable concentrations, although ratios of modeled to actual concentrations of 2 or smaller (but larger than 1) are often quoted as acceptable. Similarly, there are no thresholds separating unacceptable model error or bias from acceptable error or bias.

S4.6 Sole reliance on the highest (NAAQS) or highest of second highest (PSD) model estimated-concentrations, which are single values among available statistical summaries of modeled concentrations, impedes statistical expression (e.g., frequency distributions) of modeled concentrations and narrows the perspective of emissions impacts.

Recommendations:

The "baseline concentration" and the highest or highest, second-highest concentration, which are single values occurring during baseline years or current years, can be influenced by extreme, infrequent, transport and dispersion meteorology. These single values are extracted from large quantities of model-estimated concentrations for each year of meteorology. EPA should encourage and explore additional statistical expressions of model-estimated concentrations to broaden perspectives of emission impacts.

---

[17] "[C]onservatism only affects the decision at the margin, by deliberately preferring [better safe than sorry], from among the inevitable errors that uncertainty creates, to favor those errors which lead to *relatively* more dollars spent [e.g., on controlling emissions] ..." *Science and Judgment in Risk Assessment* by the National Research Council relating to public exposure to toxic substances. National Academy Press, 1994, page 603. "[A]ny level of conservatism (positive, zero, or negative) corresponds to some underlying attitude toward errors of overestimation and underestimation." Id., page 607.

"[D]ebate over which conservative assumption to use ... [is] a poor substitute for an effective process to ... improve regulatory decisions by reducing both uncertainties and the need for conservative assumptions." Id., page 630. "[V]alue judgments are most appropriately dealt with as part of the ... decision making phase of the overall process. ... [Decision makers] can and should override conservative default value judgements in the ... [decision] process whenever they believe it is appropriated public policy to do so. Such departures should be clearly identified as policy and not as science." Id., page 631. So scientific reasoned data inputs to models and the metrics of model performance uncertainty and accuracy analyses are needed to qualify modeled concentrations.

Advances in modeling evolve slowly. So avenues to describe, reduce and include uncertainty and inaccuracy in decision making are needed. Uncertainty analyses and bias correction of modeled concentrations should be a routine element of air quality modeling protocols, where actual concentration data are available. Bias correction of model-estimated sulfur dioxide concentrations and PSD increment consumption (section 5.8) would neutralize differences in modeled concentrations caused by differences in protocols, such as the 2002-03 NDDH and EPA Region 8 protocols.

S5. Actual concentrations obtained with monitors in the state's PSD class I areas can be used to illustrate annual trends over (or under) an estimated PSD baseline concentration.

Observations:

S5.1 The national workshop for regulatory agency modelers has been focused on modeling without monitoring data. A web search found only one Calmet–Calpuff performance analysis using actual concentrations; this analysis was conducted in Poland.

S5.2 In NSR applications, an alternative to modeling all existing sources in a modeling domain is adding worst-case model-estimated new source contributions to actual ambient concentrations.

S5.3 Federal agencies regard actual concentrations as reliable for tracking NAAQS compliance and AQRV impacts.

S5.4 Irrespective of concern as to whether actual concentrations obtained at only one monitoring site are representative of concentrations over an entire PSD class I area, the data from one site can broaden perspectives beyond model-estimated concentrations as to air quality conditions in the area.

S5.5 ***A first-time analysis of actual sulfur dioxide concentrations (sections 6.6 – 6.8)*** illustrates that trends in coal combustion and oil production data can be used for back-in-time extrapolation of 1980 actual sulfur dioxide concentrations to (a) estimate a baseline concentration and (b) examine trends in actual concentrations over an estimated baseline concentration to track increment consumption.

- ✓ The estimated sulfur dioxide 24-hour baseline concentration for the NU of TRNP is 14.7 ug/m3 and for the SU it is 5.5 ug/m3.

- ✓ Estimated changes in actual sulfur dioxide 24-hour concentrations "over the baseline concentration" were sometimes larger than the sulfur dioxide PSD class I increment of 5 ug/m3 in both units, but have been less than 5 ug/m3 since 1998.

- ✓ Estimated changes in actual sulfur dioxide 24-hour concentrations "over the baseline concentration" never exceeded the sulfur dioxide PSD class I alternate 24-hour increment of 91 ug/m3.

Recommendation:

The lower detection level of sulfur dioxide monitoring instruments has improved; an instrument is now available for detecting ambient concentrations at reliable-calibrated levels less than 1 ppb. Instruments having an improved lower detection level are needed in western North Dakota.

S6. From the beginning of the State's periodic review, EPA Region 8, with subsequent support from other federal agency modelers, has promoted consistency when emphasizing absurd modeling objectives and methods.

Observations:

S6.1 The NDDH's 1999 modeling paralleled EPA's national application of models and used overly conservative modeling inputs. This modeling provided model-estimated increases in sulfur dioxide concentrations that exceeded actual ambient sulfur dioxide concentrations. So model input data and uncertainty in model-estimated concentrations emerged as issues of concern. Bias in model-estimated concentrations does not cancel, as does the background concentration, when using modeling protocols to estimate changes in concentrations.

S6.2 EPA Region 8 states: "EPA believes that consistency in the selection *and application of models* and data bases should be sought. ... 'Consistency ensures that air quality control agencies and the general public have a common basis for estimating pollutant concentrations, assessing control strategies and specifying emission limits.' ... While consistency is the key, the Modeling Guidelines provide EPA with the authority to approve another technique if it can be demonstrated *to be more appropriate* than those recommended in the Modeling Guidelines." [18] (Here, Region 8 cites § 1.0(d) in Appendix W to 40 CFR Part 51 but does not include the sentence "Such consistency is not, however, promoted at the expense of model and data base accuracy.")

- ✓ EPA's guidance and "policies" often go beyond implementing regulations – implicitly expanding the command provisions of those regulations. [19] For example, maximum short-term emission rates and increment-affecting emission rates (which forces a paired-in-space-and-time increment consumption) are non-

[18] See Exhibit 57, page 7.

[19] See *Environmental Reporter: Analysis&Perspective – The Perils of EPA Lawmaking Through Guidance*, Vol. 31, No. 42 (10-27-00), page 2285.

regulatory applications of definitions for "actual emissions" and "baseline concentration".

✓ Region 8 (or any other interested party) has not provided persuasive reasons as to why the NDDH should not respect law and implementing regulations regarding air quality monitoring, the "baseline concentration", "actual emissions" and the sulfur dioxide "alternate increment". (See also the State Health Officer's September 8, 2003, Findings and Conclusions in Addendum E to Exhibit 158, section 11.0.)

S6.3 The State cites subsection 165(e)(3) of the CAA which states: "Any model or models designated under such regulations may be adjusted upon a determination, after notice and opportunity for public hearing, by the Administrator that such adjustment is necessary to take into account unique terrain or meteorological characteristics of an area potentially affected by emissions from a source applying for a permit required under this part." In addition, subsection 1(c) of Appendix W to 40 CFR Part 51 states: "[T]he diversity of the nation's topography and climate, and variations in source configurations and operating characteristics dictate against a strict modeling 'cookbook'."

✓ The NDDH, on behalf of the State, has an EPA-approved State Implementation Plan under the CAA. The NDDH has conducted public hearings pertaining to modeling, the MOU Protocol and other PSD program administrative issues.

✓ In 1981, a panel of modeling experts hosted by EPA noted that the federal emphasis on consistency was impeding modeling methodology improvements. So it seems now in recurring MOU Protocol issues. The State and the NDDH have completed a thorough demonstration explaining and justifying chosen modeling data bases and methods, as described in the written record and in this paper, that are as appropriate and more appropriate than federal agency modelers' preferred data bases and techniques.

✓ Region 8 (or any other interested party) has not demonstrated that the NDDH chosen data inputs for its MOU Protocol were unreasonable and arbitrary. Those choices include: (a) meteorology, (b) the background concentrations for sulfur dioxide, (c) baseline periods of major-source normal operations, (d) PSD baseline major-source sulfur dioxide emission rates, and (e) minor-source oil production sulfur dioxide emission rates. (See also the State Health Officer's September 8, 2003, Findings and Conclusions in Addendum E to Exhibit 158, section 11.0.)

S6.4 The State also cites an implementing-rule decision which states: "Congress intended that monitoring would impose a certain discipline on the use of modeling techniques ..." and "[T]hat modeling techniques be held to earth by a continual process of confirmation and reassessment, a process that enhances confidence in modeling, as a means for realistic projection of air quality." "Though EPA has authority to require methods other than monitoring in its effort to ensure that allowable increments and NAAQS are not

violated, ... it does not have authority to dispense with monitoring as at least one element of the overall enforcement effort ..." (Alabama Power Co. v. Costle, 636 F.2d 323, 372 (D.C. Cir. 1979) as cited in Exhibit 155, footnote 9.)

✓ Region 8 (or any other interested party) has not demonstrated that its use of the Calmet and Calpuff models with its choices for model control-variable input data, weather data, emission-rate data and other input data provide estimated concentrations in better agreement with the available, actual sulfur dioxide concentrations. (See also the State Health Officer's September 8, 2003, Findings and Conclusions in Addendum E to Exhibit 158, section 11.0.)

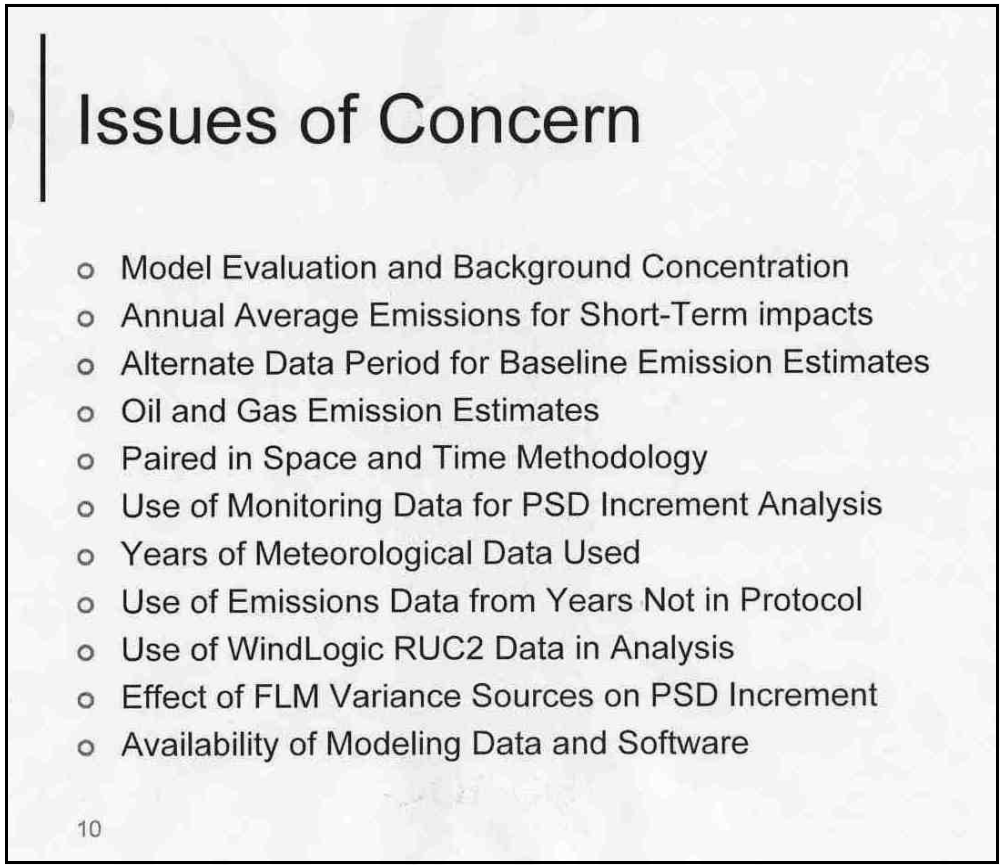
S6.5 Shadowing all modeling, subsection 165(e)(2) of the CAA states "Effective one year after August 7, 1977, the analysis required by this subsection shall include continuous air quality monitoring data gathered for purposes of determining whether emissions from such facility will exceed the maximum allowable increases or the maximum allowable concentration permitted under this part. ..."

✓ In the State's circumstance as described in S3 and S4, the technology of the Calmet and Calpuff models alone is inadequate to resolve through AQM modeled exceedance of the NAAQS and PSD increments occurring from existing or anticipated source emissions.

✓ **As illustrated in section 6.8,** a history of actual ambient sulfur dioxide concentrations obtained by monitors in TRNP blended with a longer history of coal consumption and oil production data demonstrates that current actual concentrations, due to all sources of emitted sulfur dioxide including those granted FLM certifications of no-adverse impact (CONAI) and sulfur dioxide increment variances, have not increased over an estimated baseline concentration so as to exceed the sulfur dioxide PSD 24-hour class I increment.

Table S1. A List of Recurring Issues

This list of modeling issues was presented by Kevin Golden, EPA Region 8, during a national workshop of EPA, other federal agency, state agency and local agency modelers in May 2005. This workshop was held subsequent to Region 8's participation with OAQPS in review of the NDDH's MOU modeling protocol. Patrick Hanrahan echoed similar issues, denoted with a "✓", in his August 2006 report for the Dakota Resource Council. The issues are addressed in this paper as shown below.



	Section(s) where issue is addressed.
. .	. . . . 3.3, 3.5 – 3.8, 5.1 – 5.12, 8.5
. . ✓	. . . . 4.1 – 4.3, 4.5 – 4.6, 4.10, 8.8
. . ✓	. . . . 4.4, 4.9
. . ✓	. . . . 2.2, 2.4, 4.7 – 4.9
. . ✓	. . . . 5.1, 5.3 – 5.8, 5.12, 6.1 – 6.2, 8.10
. . ✓	. . . . 2.5 – 2.8, 6.1, 6.4, 6.6 – 6.8, 8.9 – 8.11
. . ✓	. . . . 5.11
. . ✓	. . . . 5.11
. . ✓	. . . . 5.2 – 5.10, 8.5 – 8.7
. . ✓	. . . . 6.3
. . ✓	. . . . 8.6

Table S2. Glossary of Acronyms and Abbreviations

km	kilometers
lb	pound
lb/hr	pounds per hour
lb/op-hr	pounds per operating-hour
m	meters
mph	miles per hour
m/s	meters per second
mb	millibar
MMBtu	million British thermal units
ppb	parts per billion
ug/m3	micrograms per cubic meter, or $\mu g/m^3$
ADAS	ARPS Data Assimilation System
AQM	Air Quality Management
AQRV	Air Quality Related Value
CAA	Clean Air Act
CEM	Continuous Emissions Monitoring
CEMS	Continuous Emissions Monitoring System
CFR	Code of Federal Regulations
CONAI	FLM certifications of no adverse impact
DOE or DE	U.S. Department of Energy
DOI	U.S. Department of the Interior
ENSR	unknown
EPA	U.S. Environmental Protection Agency
FLAG	Federal Land Managers' Air Quality Related Values Workgroup
FLM	Federal Land Manager
FN	footnote
FR	Federal Register
FWS	DOI's Fish and Wildlife Service
H_2S	hydrogen sulfide
HSH	highest second-highest
IWAQM	Interagency Workgroup on Air Quality Modeling
LNWA	Lostwood National Wilderness Area
LRCL	lowest reliable calibrated level (in reference to monitoring)
ME	mean error

Table S2 cont. Glossary of Acronyms and Abbreviations

MB	mean bias
MNB	mean normalized bias
MM	mesoscale meteorology or meteorological model
MM4	MM version 4
MM5	MM version 5
MOU	Memorandum of Understanding
NAAQS	National Ambient Air Quality Standards
NDAC	North Dakota Administrative Code
NDCC	North Dakota Century Code
NDDH	North Dakota Department of Health
NPS	DOI's National Park Service
NSR	New Source Review
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Service
OAQPS	EPA's Office of Air Quality Planning and Standards
PCC	Pearson correlation coefficient
PSD	Prevention of Significant Deterioration
R8	EPA Region 8
RUC	Rapid Update Cycle
SH	second highest
SIP	State Implementation Plan
SO2 or SO_2	sulfur dioxide
TRNP	Theodore Roosevelt National Park
TRNP–NU	TRNP–North Unit
TRNP–SU	TRNP–South Unit
WBS	Williston Basin Study
WLI	WindLogics, Inc.

The acronyms and abbreviations listed below are described in the State's MOU Protocol which is Addendum B to Exhibit 158, page 10 through 12, or in Addendum I, table 1a.

BIAS, IPROG, R1, R2, RMAX1 and RMAX2.

Part 1

Topics Not Included

in This Paper

This paper does not provide an extensive background for the State and NDDH periodic review of the NDDH's SIP. This paper assumes that readers are familiar with details that relate to the topics listed below. Discussions of these issues are part of the State's periodic review.

Legislative history of the CAA and PSD.

The history of PSD implementing rules, including "background concentration" and "actual emissions".

The history of EPA's approved State CAA primacy.

The respective roles of EPA and the NDDH.

Details in the historical administration of PSD in North Dakota.

Details in the historical air quality modeling and model inputs under NSR in North Dakota.

The history of FLM certifications of no-adverse SO_2 impact and the NDDH's granting of SO_2 PSD increment variances.

The arbitrary and unreasonable standard.

This paper does not dwell on model estimates of sulfur dioxide PSD class I short-term increment consumption. Instead, its describes and discusses issues within Calmet and Calpuff modeling protocols which are factors affecting magnitudes of modeled sulfur dioxide concentrations.

Because the sulfur dioxide PSD class I 24-hour increment is more constraining than the 3-hour and annual increments, this paper's model performance metrics (Part 5) focus only on model-estimated 24-hour concentrations.

Part 2

Preponderant Data:

Sulfur Dioxide Emissions and

Actual Sulfur Dioxide Concentrations

2.1 Major sources of SO_2 in the state are widely scattered.

Major sources of sulfur dioxide located in the western one-half of North Dakota and in eastern Montana are shown in figure 1. The sources include coal-fired electricity generating plants, a synthetic natural gas production plant, natural gas processing plants, oil refineries and a charcoal briquette production plant.

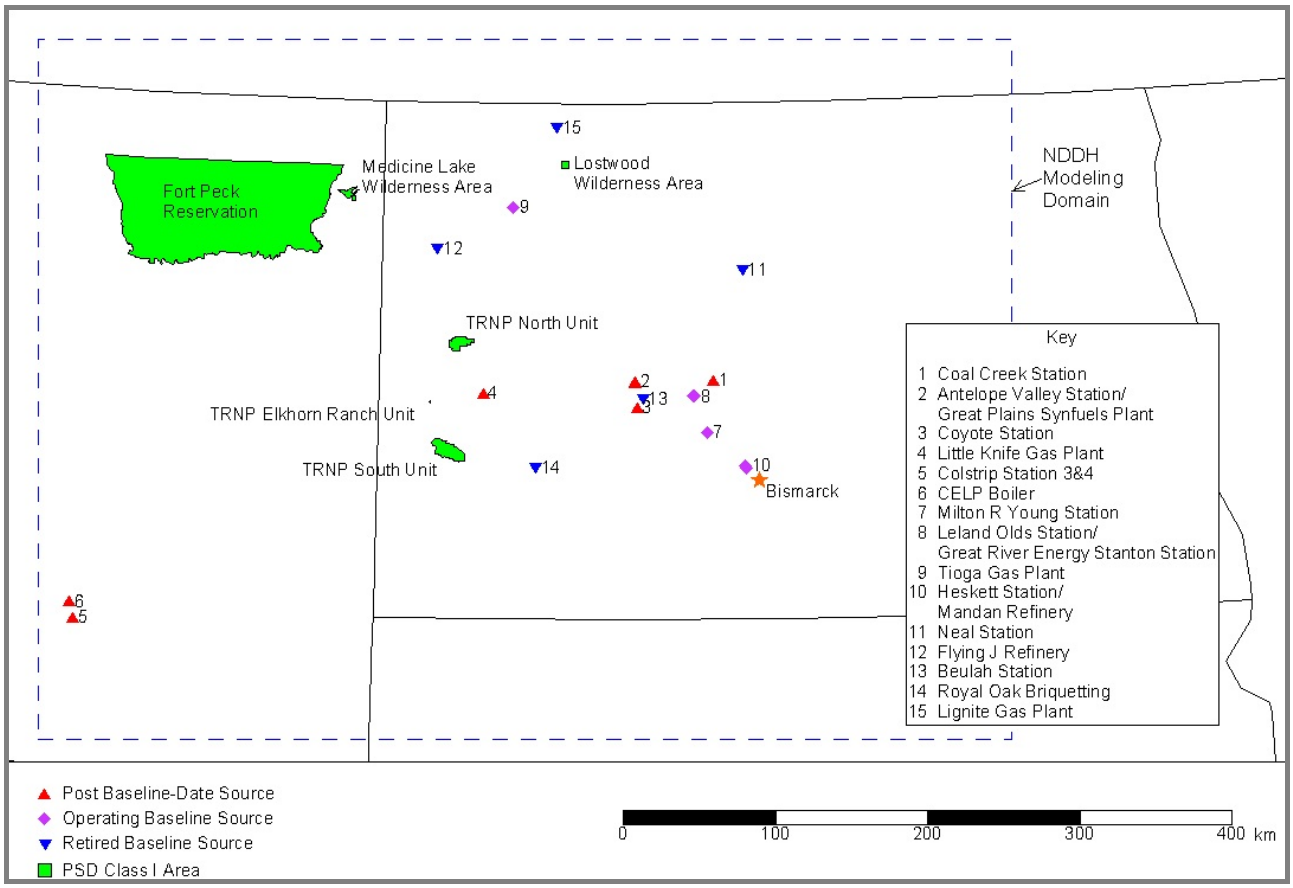


Figure 1. Map of Major Sources of Emitted Sulfur Dioxide

- Figure 1 shows locations of CAA PSD class I areas. Only an upper portion of the Lostwood National Wildlife Refuge is a CAA PSD class I area; the upper portion is known as the Lostwood National Wilderness Area (LNWA). The remaining areas of North Dakota are PSD class II areas.

- Figure 1 also shows locations of major sources of sulfur dioxide. Some major sources were constructed prior to the PSD major source baseline date (January 6, 1975); some of these sources were retired after the PSD minor source baseline date (December 19, 1977), and others are still operating. Some sources were issued a permit to construct and constructed after the PSD major source baseline date.

2.2 Annual SO_2 emissions have decreased.

The NDDH has compiled annual sulfur dioxide emissions since 1980. The sum total annual sulfur dioxide emissions for utility boilers, other point sources and oil and gas production flares and treaters are shown in figure 2.

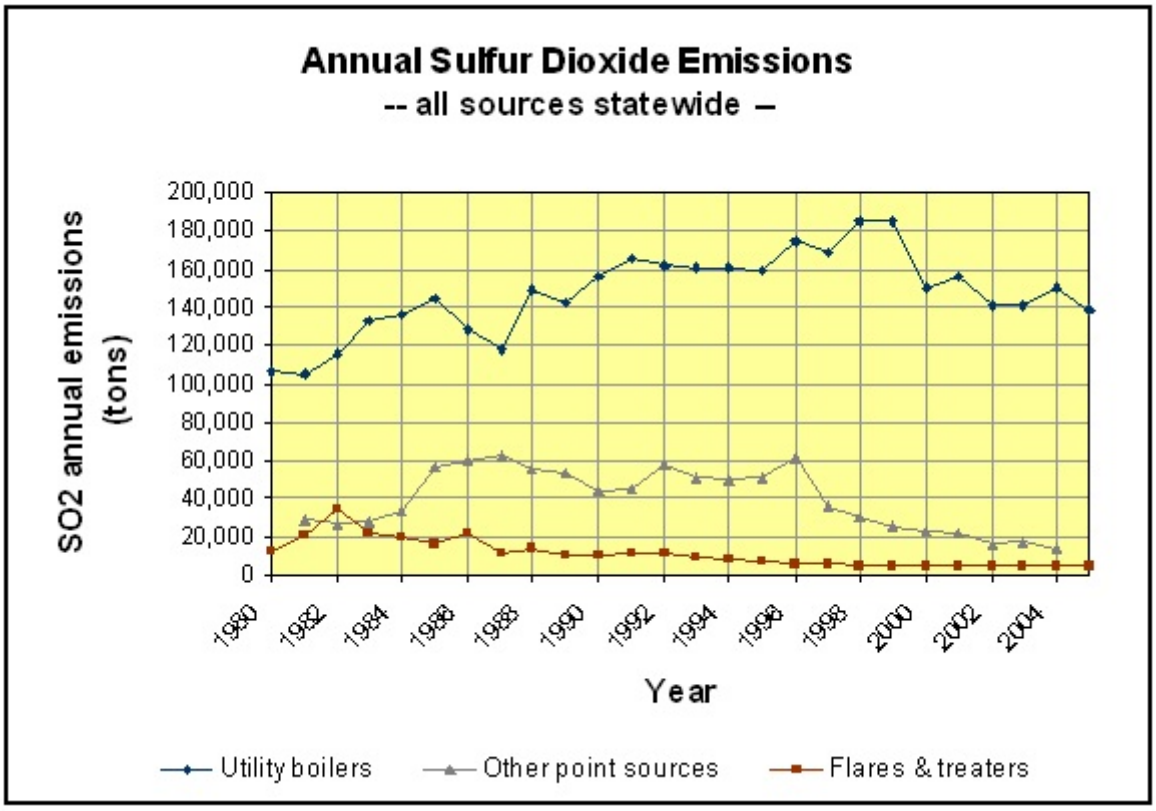


Figure 2.

- Annual utility boiler sulfur dioxide emissions peaked in 1999; since then, annual net emissions by this source category have been decreasing.

- Annual flare and treater emissions from oil and gas production sources (e.g., wells) peaked in 1982; since then, annual net emissions by this source category have been decreasing.

2.3 Utility-emitted SO_2 per unit of boiler heat input has decreased.

The annual sulfur dioxide emissions of electric utility boilers depend upon demand for power, and power production is limited by the rated maximum heat input of boilers. Increases in power production via increases in coal fired in boilers boosts annual emissions of sulfur dioxide. (See figure 2 through 1999.)

An alternate view of emitted sulfur dioxide is shown in figure 3 below.

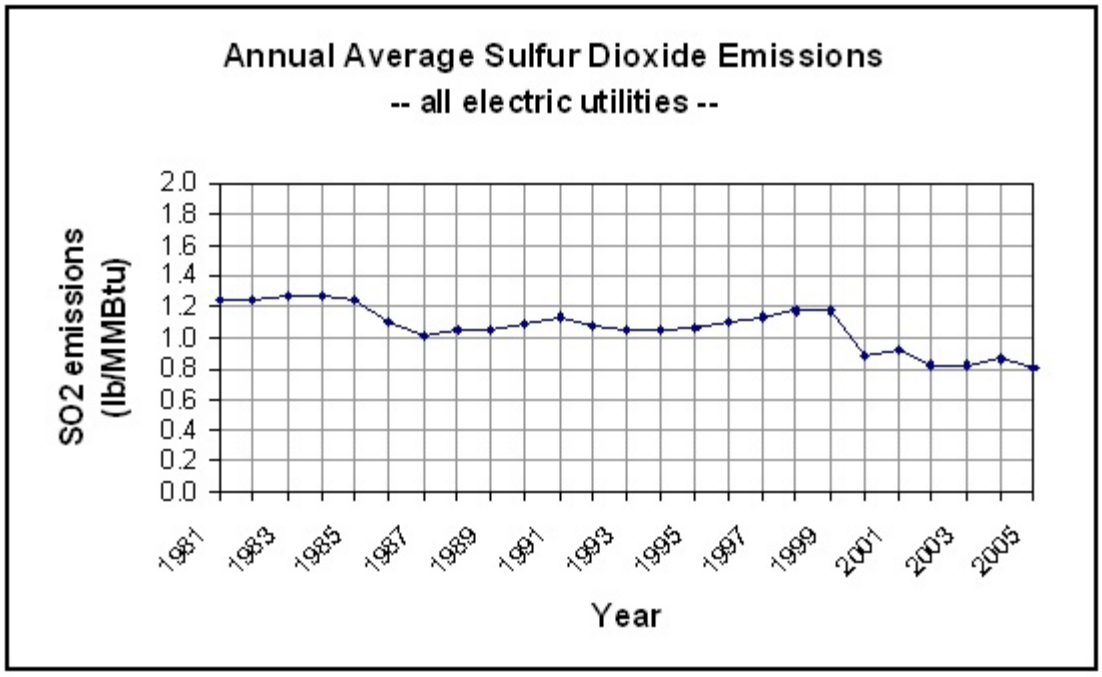
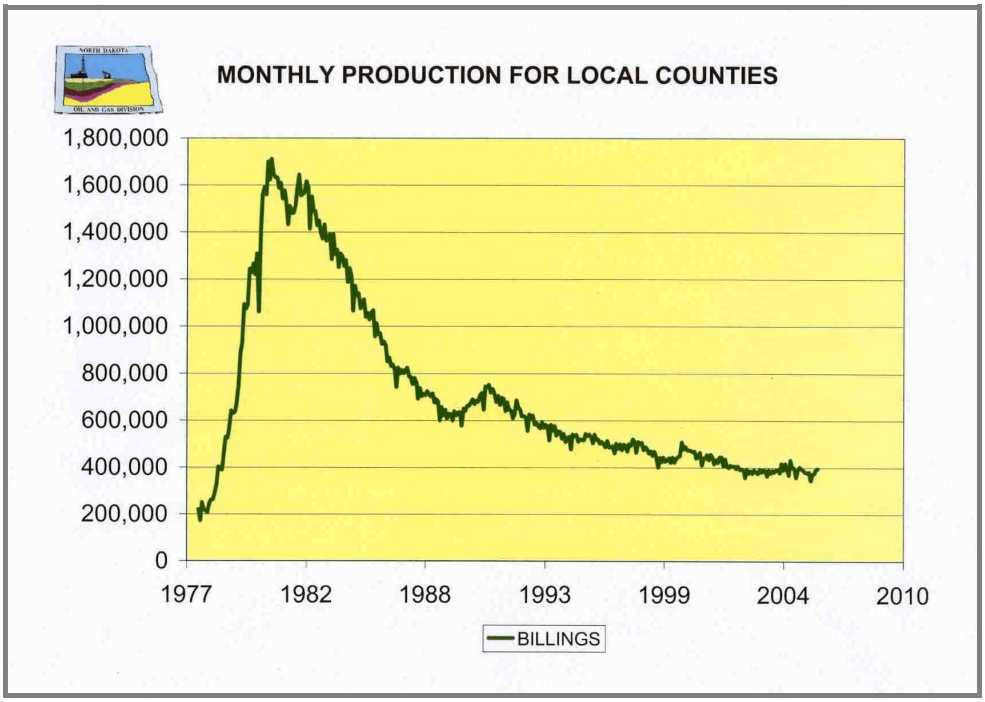


Figure 3.

- Emitted sulfur dioxide per unit of heat input to boilers (pounds per million British thermal unit or lb/MMBtu) has not increased and has trended downward (decreased) since 1999.

- The downward trend will continue after 2008 when utilities install or modify sulfur dioxide emission control systems under a CAA section 114 settlement and under CAA regional haze BART requirements.

2.4 Oil production in Billings and Dunn Counties peaked around 1982 and then decreased through 1999.



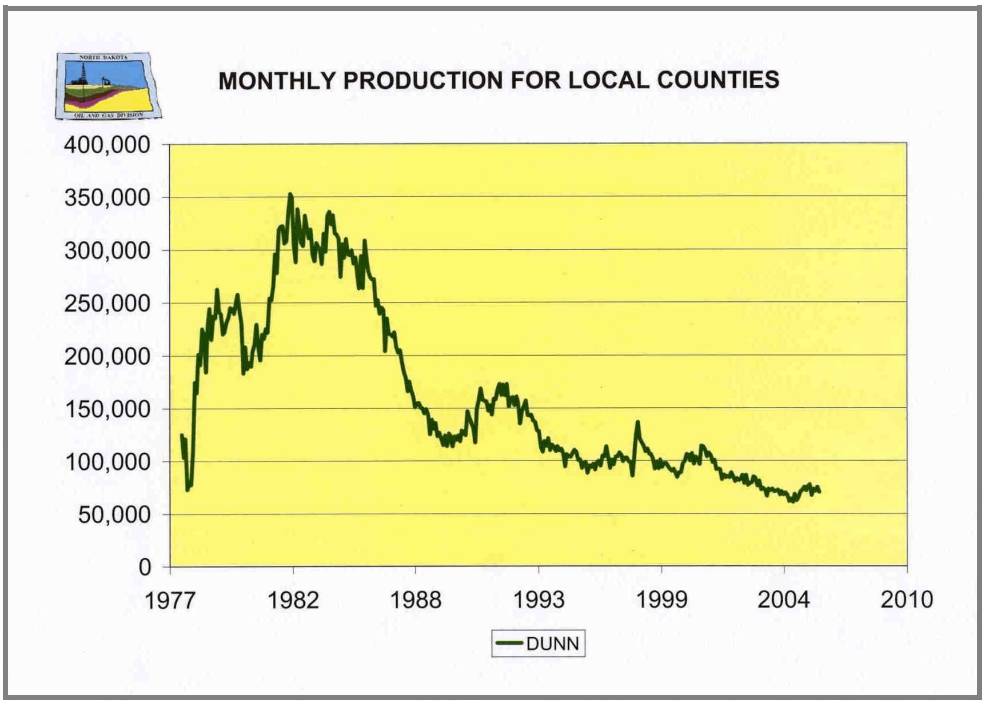


Figure 4. (Production data are annual barrels. Image copied from the web site for the Oil and Gas Division, North Dakota Industrial Commission.)

2.5 The NDDH obtains actual concentrations from several monitoring sites.

The NDDH operates several air quality and meteorological monitoring sites across the state. Air quality and meteorological data collected at these sites are used to assess trends in air pollutants such as sulfur dioxide.

Monitors at four sites provide data relevant to tacking air quality concentrations in TRNP South and North Units and to conducting model performance accuracy analyses. The four sites are TRNP–North Unit (NU) located in southeast McKenzie County, TRNP–South Unit (SU) located in central Billings County, rural Dunn Center located in central Dunn County and rural Hannover located in central Oliver County.

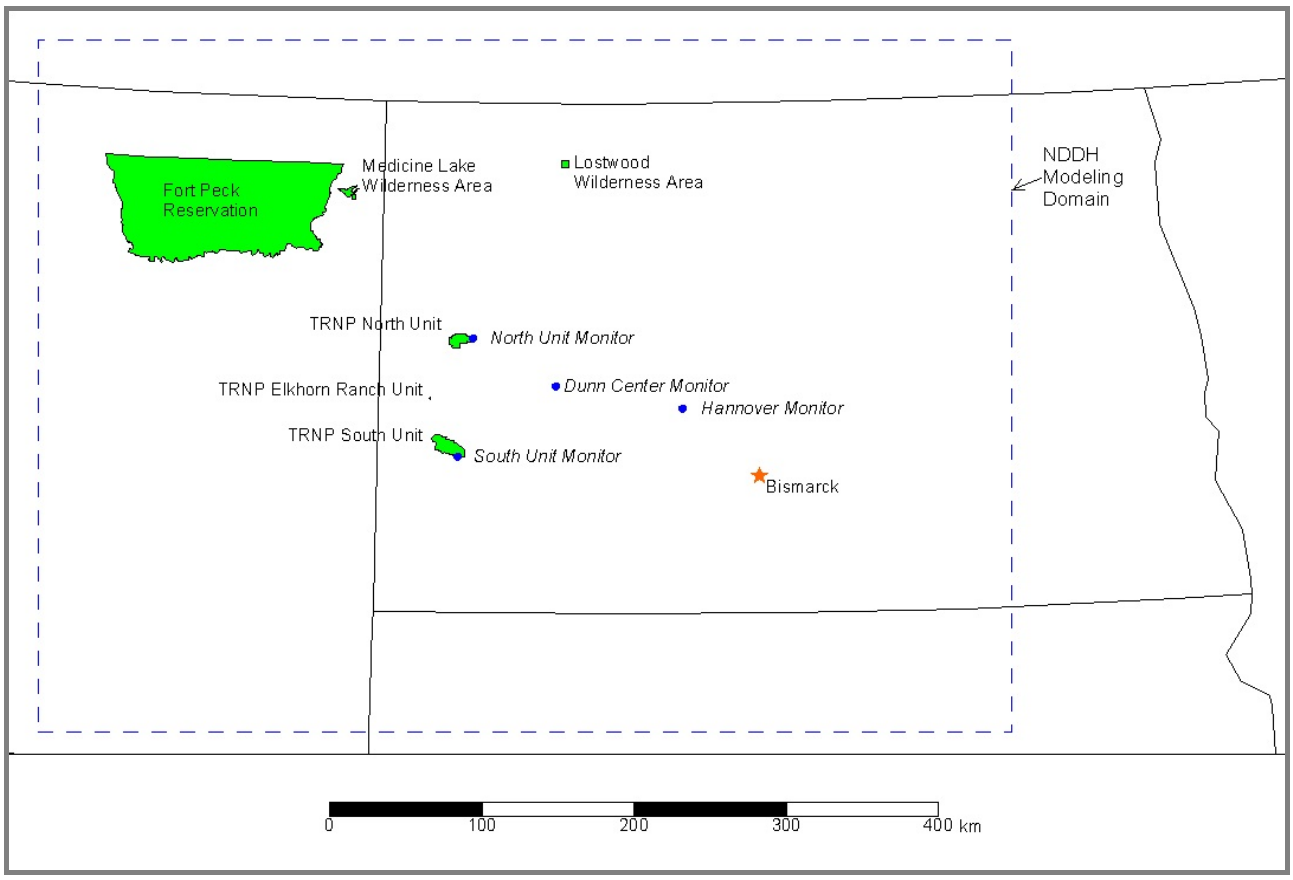


Figure 5. Map of Locations of Four Sulfur Dioxide Monitoring Sites

- In mid-1985, the monitor in the SU of TRNP was moved south to higher ground. In 2003, a monitoring site was installed at the LNWA.

- Concurrent actual concentrations, emissions data and meteorological data during one or more years are used in model accuracy tests.

2.6 Actual SO_2 concentrations are much less than EPA's human health standards.

The CAA sulfur dioxide 24-hour NAAQS for protection of human health is 365 ug/m3, not to be exceeded during any 24-hour period of the year. The 24-hour concentrations are an average of the 24, midnight to midnight, 1-hour concentrations measured with monitoring instruments.

North Dakota has always been an EPA-designated clean air state, since on-the-ground measurements of air quality have been less than all CAA NAAQS. [20] For example, the highest sulfur dioxide 24-hour concentrations have been: 162.7 ug/m3 (1982) in the TRNP-NU, 36.7 ug/m3 (1989) in the TRNP-SU, 36.7 ug/3 (1989 and 1997) at rural Dunn Center and 131.2 ug/m3 (2003) at rural Hannover.

The second-highest (SH) 24-hour sulfur dioxide concentrations for each complete year of data are shown in figure 6. Annual sulfur dioxide SH 24-hour concentrations peaked at 96.9 ug/m3 (1982) in the NU of TRNP and at 18.3 ug/m3 (1989) in the SU of TRNP. Annual sulfur dioxide SH 24-hour concentrations in both units have been 21 ug/m3 or less since 1986 and have been decreasing since 1996.

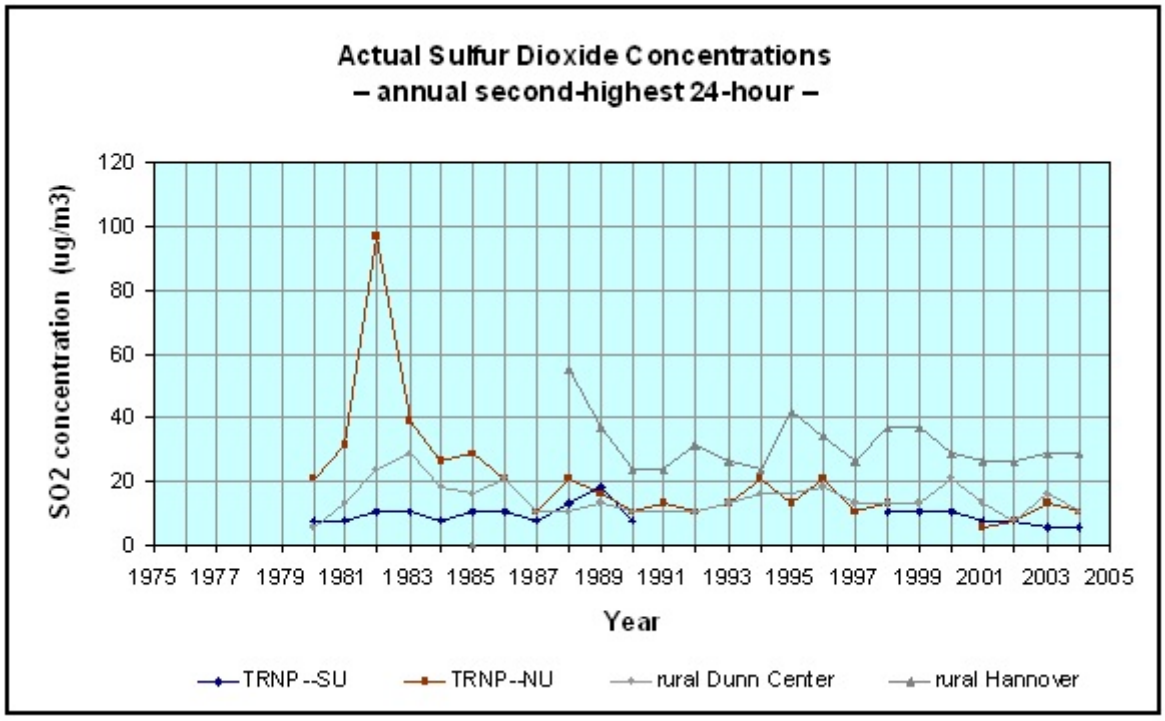


Figure 6.

[20] For detailed information, see Addendum F to Exhibit 158.

2.7 Annual occurrences of actual SO_2 concentrations at 3 ppb and larger have decreased.

The number of hours during a year when actual sulfur dioxide concentrations are detected with monitors are shown in figure 7. The threshold for this data screening is 3 ppb or 7.8 ug/m3. The data shown in the figure have not been normalized to frequencies as a percent of all hours of usable data during the year. There are 8,760 hours in one non-leap year.

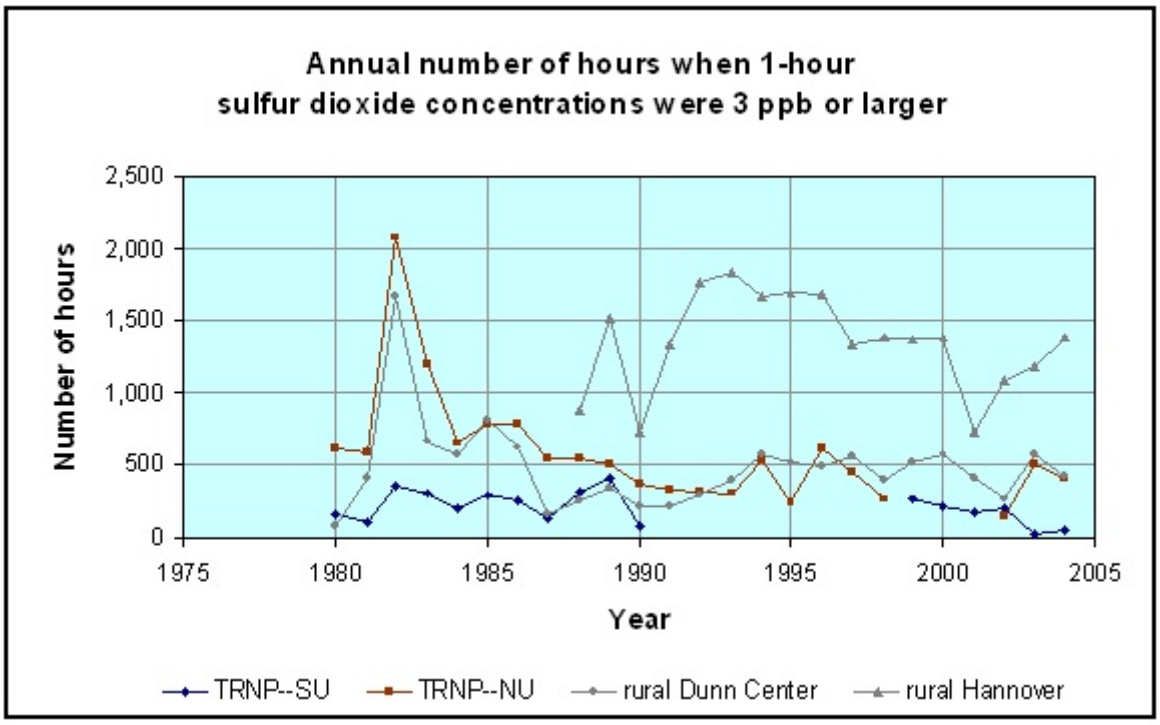


Figure 7.

The annual number of hours shown in the figure were caused primarily by (1) oil production flares and treaters and (2) major coal-fired power plants.

- (1) During the mid-1980s, the 1-hour actual sulfur dioxide concentrations at 3 ppb and larger were detected more often at the TRNP-NU and rural Dunn Center monitoring sites, which indicates detection of emissions from oil production flares and treaters. After 1986, the annual numbers of these concentrations at these sites have declined.

- (2) For years after 1987, 1-hour actual sulfur dioxide concentrations at 3 ppb and larger were detected more often at the rural Hannover site, which indicates detection of emissions from major sources in that area of the state. After 1993, the annual number of these concentrations at the rural Hannover monitoring site has declined.

2.8 Factors in addition to low wind speeds also control levels of actual SO_2 concentrations.

Daily vector averaged wind speeds were calculated from on-site 1-hour vector winds for each of the 40 days of highest daily averaged sulfur dioxide actual concentrations at the TRNP-SU monitoring site. [21] A scatter plot of the 40 daily sulfur dioxide concentrations paired with daily wind speeds for the TRNP–SU site is shown below.

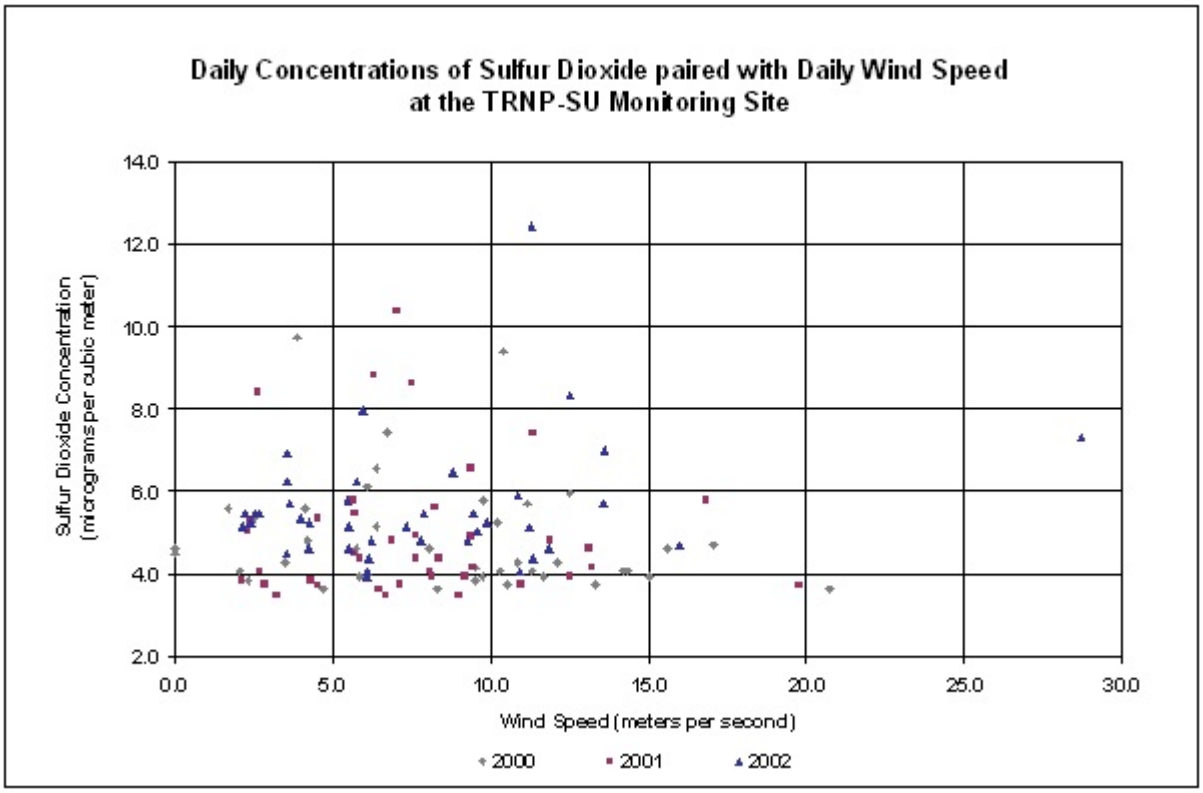


Figure 8.

- The data scatter in figure 8 illustrates that the larger 24-hour sulfur dioxide concentrations at the site of the monitor in the TRNP-SU do not always increase as wind speed decreases. (See also figures A2 and A3.)

- The reason might be due to a combination of factors, such as long transport distances between most major sources and the monitoring site, nearby minor sources, local terrain, the stochastic character of the atmospheric boundary layer and a high (1 ppb) lower detection limit for the sulfur dioxide monitor.

[21] See section 3.1 in Addendum C to Exhibit 158.

2.9 FLMs indicate no significant impact on AQRVs in the state's class I areas.

The National Park Service (NPS) assesses the status of pollutants that can impact Air Quality Related Values. [22] Its "Annual Data Summary 2005" summarizes data collected at its monitoring sites and at cooperating sites during 2005. [23] It assesses trends using monitoring data for ozone, deposition and visibility, including sulfur dioxide. [24]

Figure 9. A Vista from Buck Hill in TRNP–SU Looking to the Northeast

- The 2005 annual summary indicates that no significant environmental impacts are occurring within the TRNP or LNWA.
- The NPS conclusions are based upon data collected at one monitoring site in each of the North and South Units of TRNP.

---

[22] See Exhibit 13 in the NDDH's hearings docket.

[23] See *Annual Data Summary 2005* dated August 2006 by the NPS, report number NPS D-1782.

[24] See *2005 Annual Performance & Progress Report: Air Quality in National Parks (Draft 03/06/2006)*.

Part 3

Background for the

State's MOU Modeling Protocol

3.1 The State completed five years of legal and technical research.

- The State's process included approximately five years of painstaking legal and technical research and data analyses by State and NDDH staff and by independent legal and technical consultants.

- All modeling methods selected by the State are supported by the CAA, implementing regulations, preambles to regulations in federal registers and court decisions.

- All assembled information has been available for review and comment during the PSD periodic review process of three public hearings. Most of this information has been reviewed by independent consultants and legal experts and by some states. The information is contained in about 14,000 pages of documents, including hearing testimony and transcripts, that comprise the hearings docket.

3.2 The CAA delineates federal and state authorities.

EPA, for example, sets emission control criteria for sources and ambient air quality standards for protection of human health. The State, through the NDDH, for example, has an EPA-approved SIP and primacy under the CAA over the PSD program. On February 24, 2004, EPA and the State entered a MOU that created a process to resolve their differences in application of the Calmet and Calpuff air quality models and in tabulating model-estimated concentrations. The MOU indicates that the State has several discretionary options in modeling. [25]

✓ It may use "actual emissions" as defined by rule for 24-hour and 3-hour time periods. (See sections 4.2 and 4.6.)

✓ It may use "actual emissions" from "a different time period" other than "a two-year period which precedes" the minor source baseline date when that period "is more representative of normal source operation." (See section 4.4.)

✓ It may use emission factors based on recent continuous emission monitoring (CEM) data to estimate baseline emissions for electric utilities. (See section 4.5.)

✓ It may model current and PSD baseline emission inventories to determine estimated changes in concentrations "over the baseline concentration" and to assist in examining the correspondence between modeling and monitoring in any model accuracy analysis. (See Part 5 and sections 6.1, 6.2 and 8.8.)

The State has described discretionary options chosen and the manner in which they were applied under Section 3.0 of its SO2 PSD Air Quality Report. (See also sections in Part 4.)

[25] See Addendum A to Exhibit 158, section I.

3.3 EPA's modeling guidance does not preclude the use of a wider scope of information, data and methods.

The State's air quality modeling protocol is based on law, rule, and court decisions. The protocol is also based on EPA's discussion in preambles to implementing rules published in Federal Registers (FRs) and in guidance documents. The guidance documents are:

✓ Prevention of Significant Deterioration – Workshop Manual dated October 1980.

On page ii, EPA states: "Because this manual tends to condense the basic regulations, it may not precisely reflect the regulations and preamble thereto announced in the Federal Register on August 7, 1980 (see CFR 52676). Should there be any apparent inconsistency between this manual and the regulations published on August 7 (including any policy decisions made pursuant to those regulations), such regulations and policy decisions shall govern."

✓ New Source Review Workshop Manual – Prevention of Significant Deterioration and Nonattainment Area Permitting dated October 1990. (**Draft and never finalized.**)

In the preface, EPA states: "This document was developed for use in conjunction with *new source review* workshops and training, and to guide permitting officials in the implementation of the *new source review (NSR)* program. It is not intended to be an official statement of policy and standards and does not establish binding regulatory requirements; such requirements are contained in regulations and approved state implementation plans." (italics added)

✓ Guideline on Air Quality Models at Appendix W to 40 CFR Part 51.

In paragraph 1.a, EPA states: "The guide is not intended to be a compendium of modeling techniques. Rather, it should serve as a basis by which air quality managers, supported by sound scientific judgement, have a common measure of acceptable technical analysis."

Paragraph 1.b states: "Air quality measurements can be used in a complementary manner to dispersion models ... Measurements are particularly useful in assessing the accuracy of model estimates."

Paragraph 1.d states: "Consistency ensures that air quality control agencies and the general public have a common basis for estimating pollutant concentration ... **Such consistency is not, however, promoted at the expense of model and data base accuracy."** [26] (emphasis added)

[26] EPA Region 8 emphasizes consistency in § 1.0(d) of Appendix W, and notes that "While consistency is key, the Modeling Guidelines provide EPA with the authority to approve another technique if it can be demonstrated to be more appropriate ..." (See Exhibit 57, page 7.)

3.4 The State identified several air quality modeling principles.

Computer models have been used to estimate deterioration or improvement in air quality since 1977, a date set by the federal CAA, since no reliable on-the-ground monitoring data are available for that time period.

During the onset of its legal and technical research and inquiries, the State observed that:

- ✗ Current actual (monitored) concentrations are due to current actual emissions. And current actual 24-hour sulfur dioxide concentrations in the state's class I areas are less than 14 ug/m3.

- ✗ PSD baseline emissions were significant – about one-half of current emissions. So model estimates of baseline concentrations would be a fraction of current actual concentrations.

- ✗ Model estimates of changes in concentrations since baseline would only be similar to model estimates of current concentrations if baseline emissions were minor.

- ✗ A periodic review addresses current and baseline actual emissions, while a NSR usually addresses anticipated potential emissions of a proposed source or major modification.

- ✗ No road map existed for completion of periodic review proceedings on the status of PSD increment consumption under a SIP. **EPA, for example, has not promulgated any regulations to govern PSD programs outside a pre-construction review permitting context.**

The State adopted these principles:

- ✗ Conform actions to law, implementing rules, preambles to implementing rules, court decisions and sound engineering and earth sciences.

- ✗ Adopt or adapt federal guidance where applicable to local circumstances and identify and embrace State discretion per provisions of statute and rule.

- ✗ Act judiciously. State discretion is reasoned and justified via engineering and earth sciences disciplines.

- ✗ Ensure that actions and discretion can be documented and are transparent – subject to subsequent review and learning.

- ✗ **Conduct public hearings as required by the CAA's PSD implementing regulations for periodic review as well as by the North Dakota Century Code.**

3.5 In 1981, a panel of experts urged inclusion of model uncertainty in AQM decision-making.

EPA sponsored and hosted a workshop in May 1981 for discussion of issues relating to the use of air quality models. The findings are described in a Workshop Summary Report: Role of Atmospheric Models in Regulatory Decision-Making, EPA A-80-46, II-M-6.

> "Congress established, and required the EPA to adopt, national [ambient air quality] standards and required the states, through state implementation plans (SIPs), to develop control programs to meet the standards. This standard- (a concentration "not to be exceeded") setting strategy implicitly relied, in part, on air quality models for establishing a relationship between emissions and ambient concentration levels (air quality)." (Id., page 9.)
>
> "[T]he incorporation of such [modeling methodology] improvements has been impeded by the need for widespread consistency and standardization in the use of models in a regulatory setting." (Id., page 11.)
>
> "[F]or [Air Quality Management] AQM and the regulatory process to ignore modeling uncertainty and to continue to base decisions on best estimate single-value measures, such as the high, second-high concentrations, places an unduly heavy burden on modelers, who essentially are being required to make, or are implicitly making, policy decisions when they select models and choose model inputs." (Id., page 23.)

3.6 EPA Region 8 failed to quantify uncertainty in its modeled concentrations.

Since 1981, several EPA documents have advocated model performance accuracy analyses, which compare model-estimated and actual concentrations. See example documents below:

- ✓ Guideline on Air Quality Models, Appendix W to 40 CFR Part 51, EPA Office of Air Quality Planning and Standards, 1978 and subsequent updates. Sections 1 and 10 (now 9).
- ✓ Interim Procedures for Evaluating Air Quality Models, EPA Office of Air Quality Planning and Standards, 1981. Sections 3.1 and 4.1 and Appendix B.
- ✓ Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, $PM_{2.5}$, and Regional Haze. Draft 3.2 – September 2006, EPA Office of Air Quality Planning and Standards. Sections 16.3 and 18.4.

Region 8 failed to complete model output accuracy tests using all of its chosen Calmet and Calpuff input data, such as its larger sulfur dioxide emission rates. (See sections 3.8, 4.3, 4.5 and 4.6.)

3.7 A decision reviewing the rules implementing the 1977 PSD amendments to the CAA requires use of ambient monitoring.

"Model results performance accuracy tests have shown that the NDDH's methodology under the State's MOU protocol did not under-estimate actual sulfur dioxide in ambient air." [27]

> "The Department (NDDH) has relied on the following language for the Alabama Power case that holds that monitoring data must play the important role of holding modeling projections "to earth" through a process of confirmation and reassessment:
>
> 'We discern from the statute [28 U.S.C. § 7575(e)] a technology-forcing objective. Congress intended that monitoring would impose a certain discipline on the use of modeling techniques, which would be the principal device relied upon for the projection of the impact on air quality of emissions from a regulated source. This projects that the employment of modeling techniques be held to earth by a continual process of confirmation and reassessment, a process that enhances confidence in modeling, as a means for realistic projection of air quality. This objective is furthered by the development of sophisticated monitoring techniques, and the collection of the data base that would result from monitoring's widespread use. Of course even a congressional mandate, such as a technology-forcing requirement based on a congressional projection of emergence of technology for the future, is subject to a justified excuse from compliance where good-faith effort to comply has not been fruitful of results. That is far different from the exemption created by EPA on the basis of current technology infeasibility. Though EPA has authority to require methods other than monitoring in its effort to ensure that allowable increments and NAAQS are not violated, and though it may choose to invoke that authority because of its perception that monitoring alone is inadequate to the task, it does not have authority to dispense with monitoring as at least one element of the overall enforcement effort where Congress has mandated the use of that technique.'
>
> Alabama Power Co. v. Costle, 636 F.2d 323, 372 (D.C. Cir. 1979)."

The 2006 agenda for a national modeling workshop, which is held annually for EPA, other federal agencies, and state and local agencies, did not include discussion on the role of monitoring in modeling. [28]

---

[27] See Exhibit 155 titled *August 29, 2005 Findings and Conclusions and September 7, 2005 Determination*, footnote 9 and pages 7 and 8. See also Addendum C (MOU *Protocol Results Report*) to Exhibit 158, sections 5 and 6.

[28] See http://www.cleanairinfo.com/regionalstatelocalmodelingworkshop/

3.8 EPA Region 8 used model-results sensitivity tests when choosing inputs.

An EPA modelers' computer email memorandum dated April 21, 2004, states:

> *"Furthermore, the MOU seems to allow the use of monitoring data to calibrate modeling concentrations. According to 40 CFR part 51 Appendix W, Section 8.2.11, 'Therefore, short term model calibration is unacceptable.' "* [29] (See also sections 3.5, 3.6 and 3.7.)

In contrast, EPA Region 8 did adopt results of NDDH model sensitivity tests in construction of its 2003 draft modeling protocol, and it conducted one sensitivity test. [30] Examples follow:

- ✗ Calmet–Calpuff grid scales and vertical layers: Region 8 describes model-results sensitivity tests performed by the NDDH when selecting a grid scale and a NZ number of vertical layers. (Exhibit 84, FN 30, pages 7 and 8.)

- ✗ Calpuff switch MDISP: The Region describes results of a NDDH model-results sensitivity test when it states: "Use of dispersion coefficient option 2 provided better agreement with observations." (Id., page 17.)

- ✗ IWAQM default model-control settings: "[T]he use of IWAQM defaults increases predicted concentrations for both 3-hour and 24-hour averages. ... Had the IWAQM defaults been used in the State's limited performance evaluation, it appears that model performance would have been degraded, with the model exhibiting a bias toward over[ ]prediction." (Id., pages 15 and 17.)

Region 8 also adopted results of NDDH accuracy tests of model control settings: "To determine the effectiveness of selected Calpuff control file settings in this analysis, as well as the utility of the Calmet/Calpuff implementation in general, NDDH conducted a limited model performance evaluation, using data from two monitoring sites located in or near Theodore Roosevelt National Park. ... All seemed to have some effect on model results but, with the exception of puff splitting, none of these options caused a significant execution time penalty. Therefore, as in North Dakota's 1999 analysis, EPA has concluded it is appropriate to deploy all of these options for modeling major sources." (Id., pages 14 and 15.)

In sum, EPA Region 8 referenced the NDDH's model sensitivity and accuracy test results. Accuracy tests compare model-estimated concentrations due to current emissions to current actual observations. (Id., pages 3, 14 and 15.) Region 8 did not complete accuracy tests using its model inputs. It could not, since it only modeled changes in sulfur dioxide emission rates of major sources. (Id., pages 36 through 39.)

---

[29] See Exhibit 140, issue 1, second paragraph.

[30] See Exhibit 84. Sensitivity tests assess effects of model inputs on model-estimated concentrations.

3.9 Some EPA modelers were un-informed.

EPA regional modelers released an internet email memorandum dated April 21, 2004 [31] that comments on the initial draft of the State's MOU Protocol. [32] This memorandum was dated one week prior to EPA's verbal approval of the protocol. The modelers' remarks are in *italics* below. (See also section 7.1 and table 11.)

1. *"Even if appropriate ambient monitoring is available, it is not a sufficient replacement for modeling in assessing PSD increment consumption. Unlike modeling it cannot distinguish between sources that consume increment and those that do not, it cannot provide full spatial and temporal coverage, and it cannot provide assurance that impacts will remain low as meteorological conditions vary from year to year."*

The NDDH has not, in any protocol, replaced model-estimated concentrations with monitored actual concentrations to assess changes in air quality after PSD baseline (December 19, 1977). It has used actual concentrations to track air quality time trends, and it has compared model-estimated concentrations with actual concentrations in model performance accuracy analyses.

2 *"Variances can be provided by Federal Land Managers for sources that violate the Class I increments but do not negatively affect Air Quality Related Values. But a source with a variance still consumes increment: its emissions must therefore be included in impact analyses for other sources."*

The State has completed a detailed legal analysis of this issue. [33] The CAA and the history of PSD clearly allow deterioration in excess of PSD primary increment to levels of PSD alternate increments when Federal Land Managers (NPS and FWS) indicate that no significant adverse impacts on AQRVs are occurring. (See also section 6.3.)

3. *"Compliance of peak emissions with a 3-hour and 24-hour average increment and NAAQS cannot be assessed by averaging in lengthy periods with lower emissions. Using annual emissions for short-term impacts artificially reduces peak emissions, and is not protective of shorter-term standards."*

Modeling analyses for new source and existing source attainment of NAAQS and PSD increments are distinctly different analytical problems. Modeled worst-case concentrations using averaged emission rates, as permitted by rule defined "actual emissions", are not less than actual (monitored) worst-case concentrations. (See sections in Parts 4 and 5.)

---

[31] See Exhibit 140. See also http://www.epa.gov/Region8/foia/ndair/ndmp.html

[32] Apparently, authors of this memorandum assumed that all Region 8's concerns, dated May 24, 2002 (see Exhibit 57), regarding NDDH's May 2002 draft protocol/report also applied to the State's MOU Protocol.

[33] See Addendum H in Exhibit 158 titled *The PSD Variance Issue in North Dakota*.

4. *"Increment consumption is determined by examining the air quality degradation due to a new source, for each modeled receptor and period, not by examining the degradation relative to the day with worst existing air quality. The latter 'unpaired' procedure allows degradation of the cleanest days all the way to the condition of the dirtiest day."*

The CAA and implementing rules define "baseline concentration", which is the benchmark for assessing air quality deterioration. The NDDH used the paired-in-space-only method and EPA's paired-in-space(place)-and-time method, which ignores (since 1978) these provisions of law and rule. The CAA and implementing rules do not prohibit additional methods for tabulation of model-estimated concentrations. (See sections 6.1, 6.2, 8.9 and 8.10.)

5. *"There is no basis for averaging concentrations from different receptor locations to determine overall impact. Impacts vary in space, and averaging has the effect of artificially reducing the maximum modeled impact."*

The NDDH did average time-concurrent model-estimated concentrations among class I area receptors under its 2002 and 2003 draft protocols. However, its MOU Protocol does not use averages for receptor networks to assess air quality deterioration. Of note, FLAG lists deposition to soil and water and visibility as air quality-related values; impacts on these values are not point, but rather area or line-of-sight dependent. [34]

6. *"Baseline emissions are based on actual emissions during the baseline year, except for rare and unusual circumstances such as a labor strike or major equipment failure. Other years may need to be considered, but only insofar as they yield a better estimate of what actual emissions were at the baseline date."*

The NDDH used boiler heat input as an indicator of power demand to establish normal operations during PSD baseline. This approach is consistent with EPA preambles regarding normal operations during PSD baseline at 45 FR 52718 and during current baseline at 57 FR pages 32324 and 32325. (See section 4.4 and also sections 6.2, 8.2 and 8.8.)

– *"For large groups of minor sources, such as the hundreds of oil and gas sources in the protocol study area, it is unlikely that they will actually all be operating simultaneously. Assuming that they are would artificially inflate the baseline and available increment."*

An analysis of model-estimated concentrations without contributions of oil and gas production flares and treaters would not change the State's conclusion that short-term deterioration does not exceed PSD sulfur dioxide short-term increments. (See sections 4.7 – 4.9.)

---

[34] See Exhibit 18, which is FLAG's Phase I Report. Deposition (mass per unit area) is generally not specific to the point location of one model receptor. (Id., pages 4 and 133.) FLMs have not identified areas that might be sensitive to deposition. (Id., page 12.) Visibility of scenic vistas depends upon the color changes or light extinction (distance $^{-1}$) over the light pathway between the observer and the vista. (Id., pages 26 and 27.) Visibility also is not specific to the location of a single model receptor.

3.10 Modeling during 1999 through 2003 was draft modeling.

Historically, computer models have been required to estimate deterioration, or improvement, in air quality since 1977, a date set by the federal CAA, since no reliable on-the-ground monitoring data are available for that time period.

NDDH's draft 1999 Minnkota protocol followed pre-1999 modeling practices, even though it was not consistent with the CAA, implementing rules and preambles to implementing rules. Furthermore, this protocol did not use a minor source sulfur dioxide emissions inventory for the PSD baseline years. The draft 1999 protocol (report) was never finalized.

- In a February 2000 letter, EPA Region 8 acknowledged that the NDDH's 1999 Minnkota report was a draft. [35] And Region 8 sought more information on development of emissions inventories, including inventories for minor sources. [36]

- In a March 2001 letter, EPA Region 8 once again acknowledged that the NDDH needed to refine its 1999 modeling analysis. [37]

The NDDH provided EPA Region 8 with an initial proposed protocol on April 2, 2001. EPA Region 8 responded with criticisms on June 25, 2001. [38] Although this initial protocol was executed and a draft report prepared, [39] this protocol and the report were not topics for the NDDH's 2002 public hearing. [40]

During the years 2002 and 2001, the NDDH and EPA Region 8 did not resolve their differences on modeling issues. So during the period mid-2001 through mid-2003, each agency twice conducted separate computer modeling of sulfur dioxide emissions to assess the status of attainment of CAA PSD class I 24-hour and 3-hour increments. [41] During these years, the NDDH conducted a periodic review hearing under North Dakota law and 40 CFR § 51.166(a)(4), provided information and received written and oral comments.

---

[35] See EPA's letter dated February 1, 2000, which is part of Exhibit 17.

[36] The draft 1999 Minnkota report (which is Exhibit 129) was the first instance in modeling since the mid 1970s where the NDDH included a PSD baseline inventory of sulfur dioxide emissions for major sources. See Exhibit 82, figure 2 on page 48.

[37] See Exhibit 131, last paragraph on page 1 continuing on page 2. This document is a letter from Richard R. Long, EPA Region 8, to Francis J. Schwindt, NDDH, dated March 28, 2001.

[38] See documents in Exhibit 17.

[39] See Exhibit 7, page 5.

[40] Id, page 50. See also Exhibits 9 (Notice of Hearing) and 48 (Transcript of Hearing).

[41] The NDDH's 2002 draft report is Exhibit 6, and EPA Region 8's 2002 draft report is Exhibit 8.

- EPA Region 8's 2002 and 2003 modeling protocols (and reports) have not been issued in final form. [42] The NDDH provided EPA with comments on both protocols. [43]

When constructing the MOU Protocol, the NDDH examined EPA's 2003 protocol, assuming that this protocol reflected Region 8's preferred modeling inputs and methods. Subsequently, **the NDDH adopted all EPA (OAQPS and R8) comments on drafts of its MOU Protocol**. [44]

- EPA Region 8 participated in discussions during development of the MOU and in review of drafts of the State's MOU Protocol. During this process, Region 8 did not provide comments on sulfur dioxide emissions by oil production flares and treaters.

- The MOU Protocol differs from 2003 protocols used by the NDDH and EPA Region 8, [45] and each agency's 2003 protocol is not the same as their respective 2002 protocols.

- Neither the NDDH's 2003 protocol nor its MOU Protocol followed the industry Calmet–Calpuff protocol and analysis presented at the NDDH's May 2002 hearing. The NDDH did not use RUC data with its draft 2002 and 2003 modeling protocols, and Region 8's characterization of deficiencies in that industry protocol [46] do not apply to the NDDH's 2002, 2003 and MOU protocols.

- The major changes in successive NDDH modeling protocols from 1999 through 2003 have been described. [47] Changes in the NDDH's 2003 protocol for constructing its MOU Protocol are also described. (See section 7.1 and table 11.)

The meteorological, sulfur dioxide emissions and other Calmet and Calpuff input data used by the State in its MOU Protocol were approved by EPA (1) orally in April 2004 in advance of protocol implementation and (2) in writing in June 2005 after a preliminary review of protocol implementation.

---

[42] See Exhibit 86 (68 FR 28211). See also letter from Robert E. Roberts, Regional Administrator, EPA Region 8, to Reed Zars, attorney for the Dakota Resource Council, dated August 17, 2006, which states: "EPA does not believe the situation in North Dakota merits further EPA investigation or an independent EPA determination as to whether the SO2 increments are exceeded in the area."

[43] See Exhibits 27, 109 and 82.

[44] See Addendum I to Exhibit 158.

[45] The NDDH and Region 8 2003 draft protocols are Exhibits 81 and 84, respectively.

[46] See Exhibit 84, page 9. ENSR, the industry consultant, responded to all EPA and NDDH comments that are described on page 9 and revised its May 2002 hearing testimony. See Exhibit 95, Tabs B and C titled *Revised CALPUFF Analysis with Year 2000 MM5 Meteorological Data: PSD Increment Consumption in Class I Areas in North Dakota and Eastern Montana*, March 2003.

[47] See Addendum I to Exhibit 158, pages 2 through 4.

Part 4

Actual and Estimated

Rates for Sulfur Dioxide Emissions

4.1 When modeling for increment consumption, a current-period and a PSD baseline source inventory of emitted SO_2 are required.

Some major sources of sulfur dioxide were constructed prior to the PSD major source baseline date (January 6, 1975); some of these sources (cluster A) were retired after the PSD minor source baseline date (December 19, 1977), and others are still operating (cluster B). [48] Other major sources (cluster C) were issued a permit to construct and constructed after the PSD major source baseline data. (See also figure 1.)

Source clusters	PSD baseline sources	Current period sources
A -- Sources operated during PSD baseline and retired prior to current period.	Sources operated during PSD baseline and retired prior to current period.	
B -- Sources operated during both PSD baseline and current period timelines.	Sources that were operated during PSD baseline -- and were also operated during the current period.	Sources that were operated during current period -- and were also operated during PSD baseline.
C -- Sources constructed after PSD baseline and operated during current period.		Sources constructed after PSD baseline and operated during current period.

Cluster-A source emissions and those Cluster-B sources where emissions have decreased offset source emissions increases, whether in Cluster-C or Cluster-B.

Cluster-A source emissions and those Cluster-B sources where emissions have decreased are PSD increment-expanding sources. Cluster-C sources and those Cluster-B sources where emissions have increased are PSD increment-consuming sources.

In sum, two inventories of sources emitting sulfur dioxide are required: one for the current-period (a.k.a. current baseline or period-of-concern in NSR for NAAQS analyses or for PSD increment analyses) and another for PSD baseline.

[48] See Addendum B to Exhibit 158, Appendix C.

4.2 Modeling that uses averaged emissions during operating hours is allowed by rule. (See discretionary option 4 per MOU.)

EPA's 1978 regulatory definition for *actual emissions* was expanded by a rule change in 1980. **The rule change for *actual emissions* included the average rate, in tons per year, of a pollutant emitted during the operating hours of two consecutive years, unless those two years were not representative of source normal operations.** [49] Sulfur dioxide emission rates under this provision of the definition are reduced to pounds per operating hour (lb/op-hr).

NDDH:

The regulatory definition for *actual emissions* was applied to both periods – PSD baseline and current period (or period of concern).

- *Actual emissions*, as an average rate, is explicitly applicable to the two years of *normal operations* that represent the "period of concern" "preceding the date as of which increment consumption is being calculated." [50] In like manner, the term *actual emissions* is applicable for PSD baseline. (See sections 4.4 and 6.1.) **The NDDH determined *actual emissions* for all sources emitting sulfur dioxide, so as to provide comparable rates between both baselines and among the sources.**

- When there is no history of operation of a facility such that there are no records for emissions data, then the common practice would be to use an anticipated potential short-term pollutant emission rate. But in a periodic review, or NSR where such new source likely makes a contribution to ambient concentrations due to existing emissions of other sources, there is a history of existing emissions information.

- NDDH's sulfur dioxide emissions rates for major sources are *actual emissions* — an average of total annual emissions during all annual operating hours, not during the sum of all operating and non-operating hours (all 365 days) during the year. [51] It used these rates in its 2002, 2003 and MOU Protocol modeling for all hours throughout the year.

EPA Region 8:

The regulatory definition for *actual emissions* also includes an option for permit allowable emissions. This option would have been inappropriate, and Region 8 did not use it. (See sections 6.1, 8.2 and 8.8.) Instead, Region 8 used a non-regulatory 90th percentile of 24-hour averages of hourly CEM sulfur dioxide emissions data or a non-regulatory average of total annual sulfur dioxide emissions during all 365 days. (See section 4.6.)

---

[49] See NDAC 33-15-15-01(1)(a) or 40 CFR § 51.166(b)(13). See also 45 FR 52718.

[50] Id.

[51] See Exhibit 83, pages 37 through 89.

4.3 When modeling, averaged SO_2 rates do not introduce risk for underestimating concentrations.

Application of PSD and current baseline rates as input to Calpuff assumes that each source respectfully emits at the computed rate each hour throughout the year. EPA Region 8 states:

> "EPA found that the 90th percentile cumulative emission rate (i.e., the sum of all of the 90th percentile emission rates at each facility) did actually occur several times during the 1999-2000 time period ... In viewing the 2001 and 2002 CEM data, the results are very similar to 1999-2000. The 90th percentile of the 24-hour averages are slightly higher than the maximum cumulative total that actually occurred on any one day, however, there were numerous periods when hourly cumulative emissions exceeded the 90th percentile values used in modeling." [52]

The NDDH has also examined this matter. CEMS at power plants measure sulfur dioxide emissions each hour. Ninetieth percentiles of hourly CEMS sulfur dioxide emissions are larger than 90th percentiles of block 24-hour averages of hourly CEMS emissions, which are larger than averages of hourly CEMS emissions during operating hours. [53] Frequency distributions of hourly CEMS sulfur dioxide emission are not bell-shaped, but instead are highly skewed. [54]

- So the hourly sums of concurrent hourly CEMS emission rates for all power plant sources for each hour exceed the sum of respective source actual emission rates about 26% of all hours, rather than 50%, throughout the year. [55] (See also figure A4.)

- And hourly sums of concurrent hourly CEMS rates for all plants exceed the sums of respective source 90th percentile rates only about 1.5% of hours throughout the year.

Meteorology governs the direction of transport and the dilution and depletion of pollution during transport from sources.

- Winds transport pollutants westward from respective power plants toward respective PSD class I areas about 3 to 5% of all hours during the year as shown in figure 10. [56]

---

[52] See Exhibit 84, page 20.

[53] See Exhibit 33, table 3. This exhibit is titled *Major source sulfur dioxide data for years 2000 and 2001*.

[54] See figure 5 in Exhibit 133.

[55] See Exhibit 33, page 3. See also Appendix G to Addendum B to Exhibit 158. Addendum B is the State's MOU Protocol and is titled *A proposed alternate air quality modeling protocol to examine the status of attainment of PSD Class I increments*.

[56] See section 5.1 in Addendum B to Exhibit 158.

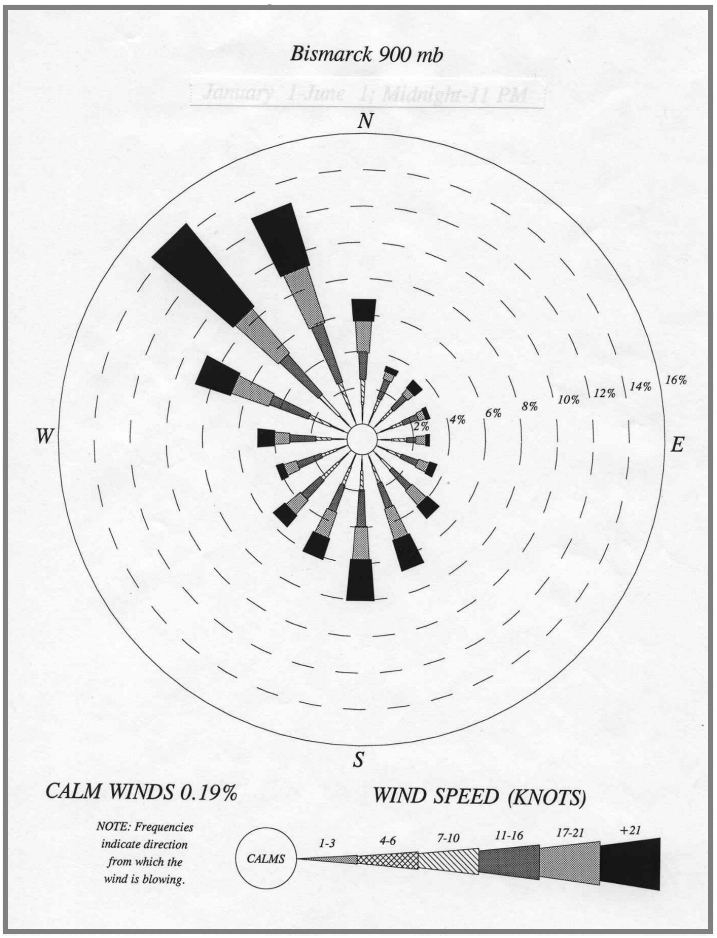


Figure 10. Wind Rose at 900 mb Pressure Altitude. Data were extracted from NWS Bismarck upper air data. The 900 mb pressure altitude corresponds to stack or plume height of most major sources of sulfur dioxide.

→ Actual and model-estimated sulfur dioxide concentrations are proportional to wind speed; concentrations increase when wind speeds decrease and decrease when wind speeds increase. As shown in figure 10, 900 mb wind speeds less than 6 knots (6.9 miles per hour) occur ~1% of time, irrespective of direction. However, surface wind speeds less than 5 mph in western North Dakota occur 14.7% of time, irrespective of direction. [57]

So the worst-case probability that models could possibly underestimate concentrations, when using rates for sulfur dioxide emissions as an annual average during operating hours, is the multiplication product of 26% times 5% times 14.7%, or 0.2%.

Likewise, the worst-case probability that models could possibly underestimate concentrations, when using 90th percentile rates, is the multiplication product of 1.5% times 5% times 14.7%, or 0.01%.

In sum, the confluence of the larger emission rates of sources, easterly or south easterly wind directions, and low wind speeds is unlikely. So model underestimates of concentrations in the state's PSD class I areas are unlikely when using *actual emissions* as an average of annual total emissions during all operating hours, or as 90th percentiles of hourly CEMS emissions.

[57] The percent-of-time number was extracted from the NDDH's 2000-02 wind data collected at air quality monitoring sites located at rural Dunn Center and in the NU and SU of TRNP.

4.4 Utility normal operations relate to demand for power. (See discretionary option 2 per MOU.)

The western North Dakota (Air Quality Control Region 172) PSD minor source baseline date is December 19, 1977. Pursuant to EPA's implementing rules and preambles to implementing rules, a PSD baseline sulfur dioxide emission rate (e.g., annual tons emitted during hours of operation) for each electric utility boiler was calculated for two consecutive years representative of source *normal operations*.

In its May 2003 draft report, [58] EPA Region 8 states:

> "The two-year study period should generally be the two years preceding the minor source baseline date, provided that the two-year period is representative of normal source operation. Another two-year period may be used, only if that other period is more typical of normal source operation than the two years immediately preceding the baseline date (see 45 FR 52718, August 7, 1980)."

Here, Region 8 paraphrases preamble text for a rule revision in the 1980 Federal Register as applying to PSD baseline. The actual text relates to the "two-year period of concern preceding the date as of which increment consumption is being calculated" (i.e., for NSR). When describing implementation of "baseline concentration", EPA states at 45 FR 52714:

> "If a source can demonstrate that its operation after the baseline date is more representative of normal source operation than its operation preceding the baseline date, the definition of actual emissions allows the reviewing authority to use the more representative period to calculate the source's actual emissions contribution to the baseline concentration."

Some coal-fired boilers had start-up operating problems. For example, Unit 2 of the Leland Olds plant had start-up problems that extended into year 1976, [59] and Unit 2 of the M.R. Young plant had start-up problems during year 1978.

An indicator of annual power plant operating activity is the amount of net power generation as kilowatt hours. These data are not available for PSD baseline years, but they are available for most units of power plants from the North Dakota Tax Department for years 1990 and later. The data can be reduced to net generation per operating hour.

An alternate indicator of activity is the annual utilization of combustion systems of power plants. Annual utilization is a ratio computed as the annual average hourly heat input of coal

---

[58] See Exhibit 84, which is EPA Region 8's *Dispersion Modeling Analysis of PSD Class I Increment Consumption in North Dakota and Eastern Montana*. May 2003. Page 30.

[59] See, for example, Exhibit 112, pages 180-184.

fired by a system's boiler divided by the boiler's rated maximum heat-input capacity. The annual average hourly heat input during system operation is the annual total heat content of coal burned during operation divided by the annual total hours of operation.

EPA addressed *normal operations* for electrical generating plants in the preamble to its July 21, 1992 revisions to PSD rules. In its analysis of comments on the proposed rule, EPA states:

> "Many commentors questioned EPA's proposed presumption that sources may use, as the baseline, emissions from any 2 consecutive years within the 5 years prior to the proposed [plant] change without regard to normal source operations.
>
> "As discussed in the proposal, this presumption is consistent with EPA's decision in WEPCO and the 5-year period for contemporaneous emissions increases and decreases in 40 CFR 52.21 (b) (3) (I) (b). Moreover, EPA is not reading normal source operations out of the regulation as charged. Rather, the presumption recognizes the nature of utility operations without compromising the existing regulatory language which requires that the pre-change 2-year period used in defining baseline emissions be representative of normal operations. For example, as a system a utility's normal operations means directly responding to a demand for electricity. A cold winter or a hot summer will result in high levels of normal operations while a relatively mild year will produce lower normal operations. By presumably allowing a utility to use any 2 consecutive years within the past 5, the rule better takes into consideration that electricity demand and resultant utility operations fluctuate in response to various factors such as annual variability in climatic or economic conditions that affect demand, or changes at other plants in the utility system that affect the dispatch of a particular plant."
>
> See 57 FR, No. 140, pages 32324 and 32325. [60]

A five-year period was used in like manner per EPA interpretation of rule regarding current baseline. Data for computation of the utilization of the rated heat-input capacity of each coal-fired combustion system of power plants for each year from 1975 through 1980 were assembled. [61]

In sum, the two consecutive years having the greatest power demand, as reflected by greatest use of heat-input capacity of the plant's systems, were chosen as representative of source *normal operations*. Coal sulfur-content data were not used or considered in selecting the two consecutive years of source normal operations.

---

[60] See also Exhibit 83 titled *May 2003 Prevention of Significant Deterioration Sulfur Dioxide Final Baseline Emission Rates with Appendices A through K*, pages 13 and 14. And see also 40 CFR § 51.166(b)(47), 40 CFR § 52.21(b)(48) and Appendix D to Addendum B in Exhibit 158.

[61] See, for example, section 4.0 and Appendix D of Addendum B to Exhibit 158.

4.5 EPA's AP-42 emission factor for SO_2 underestimates PSD baseline emissions. (See discretionary option 3 per MOU.)

When no CEMS data are available, annual sulfur dioxide emission rates (such as pounds per hour) are calculated as the total sulfur dioxide emissions during all hours of operation.

Annual uncontrolled sulfur dioxide emissions for boilers of electric power plants are calculated as the product of multiplying the annual coal fired in boilers (tons or pounds) by the average sulfur content of the coal and an emission factor (since some sulfur is retained in ash or slag). The emission factor depends on the amount and form of natural scrubbing agents (primarily sodium oxide) in coal, boiler design and boiler load.

PSD baseline emission factors were assumed to be the same as current-period emission factors. A current-period sulfur dioxide emission factor was calculated for each of five boilers of power plants that did not have sulfur dioxide emission controls during current and PSD baselines. [62] Current-period source sulfur dioxide emission factors are a sulfur-balanced calculation comparing annual sulfur dioxide emissions using recent CEMS data with corresponding coal consumption and coal sulfur-content data. The outcome of these calculations indicates that emission factor multipliers for the five boilers range from 27.0 to 38.7.

The NDDH's approach, which computed emission factors for each of the five electric utility boilers that were operated during the current period and PSD baseline, provides an "apples-to-apples" comparability between current period and PSD baseline sulfur dioxide emission rates.

The multiplier used by Region 8 was 30 [63] per EPA's AP-42. [64] Multipliers in emission factors depend on the amount of sodium oxide, and to lessor extent other alkali constituents, in reactive form in coal combustion ash. [65] In general, multipliers increase as sodium oxide increases in combustion ash. For example, a multiplier of 30 is equivalent to a 25% retention of sulfur during coal combustion in boiler ash and slag, while a multiplier of 40 is equivalent to 0% retention. The sodium oxide in lignite combustion ash varies in the range of 2 to 8%, [66]

---

[62] See Exhibit 83, pages 37 through 89.

[63] See Exhibit 84, pages 31, 32 and 33.

[64] See Exhibit 120, EPA's *Emission Factor Documentation for AP-42 Section 1.7, Lignite Combustion*, Acurex Environmental Corporation, RTP, NC. April 1993.

[65] See Appendix E in Addendum B to Exhibit 158, pages 42 – 44.

[66] Id. See also: Exhibit 83, pages 25 and 30 – 32; "Analyses of the Northern Great Plans Province Lignites and Subbituminous Coals and Their Ash" by the Grand Forks Energy Technology Center dated July 1981, document number DOE/GFETC/RI-81/2 (DE81028366); and miscellaneous utility boiler coal-ash data provided by EERC, Grand Forks, per request of the NDDH.

except sodium oxide in Heskett's combustion as is less than 2%. So the AP-42 emission factor for North Dakota lignite fired boilers is 32.3S reflecting sodium oxide concentrations in the lignite. (Exhibit 120, FN 64, table 4-1.)

EPA Region 8's approach under or overestimated PSD baseline sulfur dioxide emissions as shown in table 1. If EPA Region 8 had used NDDH's source-specific sulfur dioxide emission factors, its net PSD baseline sulfur dioxide emission rates would have been larger by 1,918.2 pounds per hour, and its net increment-affecting rates would have been lower by the same amount.

Table 1. EPA R8 Boiler Baseline SO2 Emission Rates Using NDDH's Emission Factors

	SO2 Emission Factors				SO2 Emissions (lb/hr)	
Boiler	R 8	NDDH		ratio	R 8	R8 x ratio
Neal Station	30.0	32.9	a	1.10	297.0	325.7
Beulah Station	30.0	27.0	b	0.90	621.0	558.9
Hesket Unit 1	30.0	27.0	c	0.90	589.0	530.1
Leland Olds Unit 1	30.0	37.4	c	1.25	2,714.0	3,383.5
Leland Olds Unit 2	30.0	38.7	c	1.29	1,625.0	2,096.3
M.R. Young Unit 1	30.0	33.3	c	1.11	3,972.0	4,408.9
Stanton Unit 1	30.0	35.5	c	1.18	2,359.0	2,791.5
Sum					12,177.0	14,094.8

a = emission factor per stack sampling (see Addendum B to Exhibit 158, page 46).
b = emission factor set the same as Hesket Unit 1 because boilers were similar design.
c = baseline emission factors set the same as current emission factors.

Uncertainty and inaccuracy in calculation of emission factors can result from error in CEMS data, amount of coal fired in boilers and coal sulfur-content data.

In sum, the NDDH justified a deviation from EPA's AP-42 sulfur dioxide emission multiplier for several of the state's lignite-fired utility boilers. Source-specific sulfur dioxide emission multipliers, among other factors, explain differences in respective NDDH and EPA Region 8-modeled sulfur dioxide emission rates and model-estimated sulfur dioxide deterioration.

(Note: The two-year-averaged emission rate during baseline *normal operations* was computed as the dividend resulting from dividing the two-year total sulfur dioxide emissions (pounds) by the two-year total hours of operation.)

4.6 EPA Region 8 matched and mixed methods for calculating SO_2 emission rates.

Sulfur dioxide emission rates used by the NDDH in its MOU Protocol and by EPA Region 8 in its 2003 draft report are shown in table 2.

When, in 1999, the NDDH asked EPA Region 8 for guidance on the appropriate emission rate for PSD increment-expanding sources (those operating during PSD baseline but retired (shut down) prior to the current period as shown in section 4.1), Region 8 responded as follows. "As you requested, in preparing our response we have discussed the issue with OAQPS and the other Regional Offices to ensure national consistency in this issue. ... [I]t appears that your situation is a special case because of the high variability of emissions from some of the largest major sources being modeled. ... The difficulty in this case in finding a conservative method to characterize increment expanding emissions that will not result in too much 'increment credit'. *We are concerned that using the peak observed short term emission rate for each and every 3 and 24 hour period simulated in the model would overestimate increment expansion because it is extremely unlikely that the source was operating at this peak level at the time this 'worst case' meteorology actually occurred.* We believe that our suggestion to use an annual average operating rate for all averaging periods is a reasonable screening method that is equitable to all sources and will not overstate increment expansion." (See Exhibit 128, which a letter by Richard Long, Director Air and Radiation Programs, dated June 1, 1999. Italics added.)

Similarly, use of a peak short-term rate, such as a 90th percentile rate, for each source constructed after PSD baseline and operating during the current period overestimates increment consumption for the same reason; that is, it is extremely unlikely any of these sources operate at peak level at the time of worst-case meteorology. (See section 4.3.)

Current (2000-2001) rates:

- As illustrated in figure A4, EPA Region 8's current period rates are non-regulatory 90th percentiles of 24-hour averages of CEMS hourly emissions for coal-fired electric utility boilers. [67] As shown in table 2, some boilers were constructed and placed into operation after the PSD minor source baseline date. Region 8 used non-regulatory average annual rates for an oil refinery, natural gas processing plants and a synfuels plant.

[67] Region 8's choice of 90th percentile rates rather than maximum emission limits departs from tables 9-1 and 9-2 in Appendix W to 40 CFR Part 51 (which is Exhibit 132) and from table C-5 and pages C.48-C.49 in EPA's **draft** *New Source Review Workshop Manual – Prevention of Significant Deterioration and Nonattainment Area Permitting* dated October 1990. See Exhibit 84, page 20. Apparently, tables 9-1 and C-5 apply to SIP compliance evaluations due to similar titles, etc., and table 9-2 applies to a PSD and NAAQS NSR. (Exhibit 132, section 9.1.2.i.) But all three tables are inconsistent with 45 FR 52714 and 52718 and rule-defined "actual emissions".

- EPA Region 8 did not have hourly CEMS sulfur dioxide emissions for the sour natural gas processing plants, the Great Plains Synfuels plant and Heskett Station Unit 1. It did not use, but had in its Acid Rain Program data base, hourly CEMS emissions for the plants in southeastern Montana. The NDDH obtained current-period CEMS hourly sulfur dioxide emissions from these sources or from the Acid Rain Program data base.

- EPA Region 8 included a sulfur dioxide emission rate for the main stack and for the bypass stack at the Great Plains Synfuels Plant. As shown in table 2, Region 8's sulfur dioxide emission rate for the bypass stack was 855 lb/hr. But the bypass stack is used only when the main stack malfunctions. So inclusion of emissions of both stacks in their modeling is a double counting and increased sulfur dioxide increment-affecting emissions by 855 lb/hr.

PSD baseline rates:

- No CEMS were available for measuring the sulfur dioxide emissions of electric utility coal-fired boilers during the PSD baseline period. So EPA Region 8 computed a ratio of 90th percentile rate from current CEMS data to average annual emissions and then used this ratio as a multiplier with PSD baseline average annual emissions (see section 4.5) to estimate PSD baseline 90th percentile rates for most utility boilers.

  But rated maximum heat-input capacities limit amount of coal fired in boilers. Boilers were likely operated at capacity sometime during PSD and current baselines; so 100th percentiles would be the same, unless other factors such as coal quality changed. [68] When this is true and when general utilization of boilers changes, the ratio of 90th percentiles to average annual emissions would not be the same for both time lines.

- However, EPA Region 8's PSD baseline rates for a charcoal briquette production plant, two electric utility plants and two oil refineries are average annual rates.

In sum, EPA Region 8 choices for sulfur dioxide emission rates are unreasonable in three respects. First, use of 90th percentile rates as surrogate peak rates for utility boilers during the current period overestimates current actual concentrations because it is extremely unlikely any of the sources operate at peak level at the time of worst-case meteorology. Second, use of 30S rather than source sulfur-balanced emission factors when calculating estimates for boiler PSD baseline rates and, third, (a) use of 90th percentile rates for boilers constructed after the PSD minor source baseline date and (b) annual average (over 365 days) rates for boilers retired (shut down) after that baseline date are not "apples-to-apples" comparable applications of emission rates; both practices distort increment-affecting emissions by widening the difference between current and PSD baseline emissions and overestimate increment consumption.

---

[68] When modeling, use of 100th percentile rates, as well as permit-allowed maximum rates per an option under the rule-defined "actual emissions", for both the PSD baseline and the current period would result in no net change in estimated concentrations, creating "the Gulf Coast problem" (45 FR 52681).

Table 2. NDDH and EPA Region 8 PSD Baseline (bl) and Current-period Sulfur Dioxide Emission Rates

			State MOU protocol lb/op-hr			EPA R8 2003 lb/hr		
	Source	Unit	PSD bl	current	difference	PSD bl	current	difference
Sources operating at PSD baseline and retired prior to current period.	Beulah Power Plant	1 &2	127.0		-127.0	with below		
		3, 4 & 5	203.6		-203.6	621.0		-621.0
	Neal Station	1 &2	364.6		-364.6	297.0		-297.0
	Royal Oak Briquetting	Boilers 1, 2 &3	220.8		-220.8	with below		
	Plant	Carbonizer Furnaces	1,124.8		-1,124.8	545.0		-545.0
	Williston Refinery	All units	51.7		-51.7	45.0		-45.0
		Preflash Heater	7.1		-7.1			
		Crude Heater	7.7		-7.7			
		Thermal Cr. Heater	0.3		-0.3			
		Charge Heater	0.1		-0.1			
		Reformer Heater	0.5		-0.5			
		Boiler 1	10.5		-10.5			
		Boiler 2	10.5		-10.5			
		Boiler 3	15.0		-15.0			
	category subtotal		2,144.2		-2,144.2	1,508.0		-1,508.0
Sources operating during PSD baseline and also operating during current period.	R.M. Heskett Station	1	415.8	248.0	-167.8	589.0	342.0	-247.0
		2	969.9	612.7	-357.2	1,625.0	849.0	-776.0
	Leland Olds Station	1	3,609.8	4,179.2	569.4	2,714.0	5,085.0	2,371.0
		2	7,312.4	8,145.1	832.7	4,185.0	10,354.0	6,169.0
	M.R. Young Station	1	4,357.0	5,161.4	804.4	3,972.0	6,087.0	2,115.0
		2	4,726.5	4,353.2	-373.3	5,635.0	5,749.0	114.0
	Stanton Station	1 (PSD baslein only) and 10	2,271.3	2,389.8	118.5	2,359.0	2,985.0	626.0
	Tioga Gas Plant	SRU Incinerator	1,107.1	300.6	-806.5	1,107.0	305.0	-802.0
	Lignite Gas Plant	SRU Incinerator	285.8	105.6	-180.2	286.0	120.0	-166.0
	Mandan Refinery	Boilers + Crude Furnace	1,172.7	133.0	-1,039.7	1,062.0	131.0	-931.0
		FCCU	1,135.8	1,026.9	-108.9	1,136.0	1,006.0	-130.0
		Alkylation Unit	160.3	7.7	-152.6	160.0	8.0	-152.0
		Ultraformer Furnaces	15.3	15.9	0.6	14.0	16.0	2.0
		SRU Incinerator		45.3	45.3		49.0	49.0
	category subtotal		27,539.7	26,724.4	-815.3	24,844.0	33,086.0	8,242.0

Table 2 cont. NDDH and EPA Region 8 PSD Baseline (bl) and Current-period Sulfur Dioxide Emission Rates

Sources constructed after PSD baseline and operating during current period.	Coal Creek Station	1		3,368.1	3,368.1		4,269.0	4,269.0	
		2		2,972.8	2,972.8		3,429.0	3,429.0	
	Antelope Valley Station	1		1,590.8	1,590.8		3,440.0	3,440.0	
		2		1,496.0	1,496.0				
	Coyote Station	1		3,955.4	3,955.4		4,755.0	4,755.0	
	Grasslands Gas Plant	SRU Incinerator		113.4	113.4		263.0	263.0	
	Little Knife Gas Plant	SRU Incinerator		80.1	80.1		84.0	84.0	
	Great Plains Synfuels	Main stack		1,094.4	1,094.4		1,051.0	1,051.0	
		Bypass stack *		none	none		855.0	855.0	
		Start-up flare		119.0	119.0		396.0	396.0	
		Main flare		184.0	184.0		178.0	178.0	
		Back-up flare		78.0	78.0		124.0	124.0	
	category subtotal			15,052.0	15,052.0		18,844.0	18,844.0	
	PPL Corp Colstrip	3		742.9	742.9		655.0	655.0	
		4		719.0	719.0		597.0	597.0	
	CELP Colstrip			419.8	419.8		420.0	420.0	
	category subtotal			1,881.7	1,881.7		1,672.0	1,672.0	
	Total all rates =		29,683.9	43,658.2	13,974.3	26,352.0	53,602.0	27,250.0	

SO2 rates are "actual emissions" as tons per year during operating hours.
SO2 rates are tons per year during all 365 days.
SO2 rates are 90th percentiles of 24-hour averages of hourly CEM data (current) or estimates of 90th percentiles (PSD bl).

In sum, a source's increment-affecting sulfur dioxide emission rate is the difference between current and PSD baseline rates. EPA Region 8's net increment-affecting sulfur dioxide emission rate is 27,250.0 lb/hr and the State's rate is 13,973.3 lb/op-hr (modeled as lb/hr). The difference of 13,276.7 lb/hr is due primarily to use of a 90th percentile of 24-hour averaged hourly CEM emissions for PSD sources (those operating during years 2000-01 but not during PSD baseline) and to much larger increment-affecting rates for the Leland Olds Station and the M.R. Young Station.

The State's current sulfur dioxide emission rates more accurately capture current ambient air quality conditions as confirmed by current monitoring data. (See Part 5 and sections 8.8, 8.9 and 8.10.)

4.7 Omission of minor sources would not change status of PSD SO_2 increment consumption.

In May 2003, EPA Region 8 states:

> "EPA believes that the State's preliminary oil and gas inventory significantly overstates the level of SO2 emissions in the 1976–1977 baseline period." [69]

> "Overestimation of the base year emissions would create more increment expansion credit than is justified, hence the total amount of increment consumption calculated in this study may be too low." [70]

Sulfur dioxide emissions of oil and gas flares and treaters during the early and mid 1980s likely did have significant impact on on-the-ground actual concentrations over the TRNP-NU area and the rural Dunn Center area and lesser impact over the TRNP-SU area. (Compare figures 2, 4 and A1 to figures 6 and 7.)

Any impact of sulfur dioxide emissions of flares and treaters cannot be statistically isolated from other source impacts via analysis of available actual concentrations and meteorological data. This statistical handicap is likely due, in part, to the 1 ppb lower sulfur dioxide detection level (since 1997) of monitoring instruments used to measure actual concentrations. A lower detection level is needed, because actual 1-hour concentrations in TRNP and at rural Dunn Center are usually less than the instrument's detection level – in TRNP, for example, more than 80% of all hours during the year.

Model-estimated sulfur dioxide deterioration was recalculated by omission of the baseline and current inventories for oil and gas sources and by using EPA's paired-in-time-and-space method. The highest second-highest (HSH) model-estimated 24-hour increment consumption decreased for the TRNP-SU by 0.27 ug/m3 to 4.37 ug/m3, increased for TRNP-NU by 0.34 ug/m3 to 4.80 ug/m3, increased for the TRNP-Elkhorn Ranch by 1.70 ug/m3 to 2.47 ug/m3 and did not change for the LNWA. [71]

In sum, a model results sensitivity test was completed without baseline and current inventories of sulfur dioxide emissions from oil and gas flares and treaters. The outcome of this test demonstrates that omission of these sources' emissions did not result in model estimates of sulfur dioxide 24-hour deterioration, since PSD baseline, that are larger than the PSD class I 24-hour increment.

---

[69] See Exhibit 84, page 21.

[70] Id., page 21

[71] See Exhibit 153 titled *Sensitivity Analysis of North Dakota MOU Modeling of SO2 Increment Consumption* by ENSR Corporation, section 5 and table 5-2.

4.8 The NDDH used average rates, not peak rates, for SO_2 emissions of oil production flares and treaters.

In May 2003, EPA Region 8 also states:

> "One major concern with the oil and gas inventory was that the estimates were based on the average of peak short-term emission rates, rather than annual average emission rates. This is a problem in estimating emissions from oil and gas sources because sources may only operate for a period of weeks or months at a time ..." [72]

Nevertheless, Region 8 has not provided details that would illustrate an overstatement of sulfur dioxide emissions from oil production flares and treaters during PSD baseline. And it incorporated the NDDH oil and gas inventories into its draft 2003 modeling without proposing or applying any adjustments to flare and treater sulfur dioxide emissions. [73]

The NDDH provided descriptions of PSD baseline and current (year 2000) inventories of oil production flares and treaters and confounding factors in calculating baseline sulfur dioxide emission rates. [74] Only sources within 50 km of the state's PSD class I areas were included in inventories. The NDDH also provided total rates for all sources and maps for these sources. [75]

The numbers for oil production sources in inventories within 50 km of the NU and SU of TRNP are similar as shown in table 3. On average, sulfur dioxide emission rates were significantly larger during PSD baseline for oil production sources near the Elkhorn Ranch and NU of TRNP compared to the average rate for sources near the SU of TRNP. Reasons for these larger average rates include larger sour-natural gas to oil ratios and larger hydrogen sulfide (H_2S) concentrations in the raw gas (before venting or flaring) [76] during initial oil production. Note, for example, that oil production during the 1980s surrounding the NU of TRNP was similar to production surrounding the SU. (Compare figures 4 and A1.) By year 2000, the field average raw gas H_2S was 2.42% for producing wells surrounding the TRNP-NU and 2.33% for producing wells surrounding the TRNP-SU. (As shown in figures 6 and 7, trends for actual sulfur dioxide concentrations at sites of monitors near Dunn Center and in the NU of TRNP converge to the nearly-flat trend at the site of the monitor in the SU of TRNP by 1989-90. See also Attachment C.)

---

[72] Exhibit 84, page 21. See also Exhibit 57, pages 9 and 10.

[73] Id. See also page 22.

[74] See Appendix F to Addendum B to Exhibit 158. See also Exhibit 83, pages 90 through 102.

[75] Id., pages 53 and 54.

[76] During initial well production, H_2S concentrations in raw sour-natural gas in the Williston Basin ranged from near 0% to more than 30%. See *Final Report. Sulfur Dioxide Emissions Inventory for Sources near the Theodore Roosevelt National Park.* February 1983 by the NDDH, page 7.

Table 4. Statistical Summary of Oil Production Sources Loacted Within 50 km of PSD Class I Areas *

	PSD Baseline Period			Current Period (year 2000)		
Class I Area	Number of Sources	Total SO2 Emission Rate (lb/hr)	Average SO2 Rate (lb/hr)	Number of Sources	Total SO2 Emission Rate (lb/hr)	Average SO2 Rate (lb/hr)
TRNP - South	196	824.6	4.21	262	318.3	1.21
TRNP - Elkhorn	248	8,030.1	32.38	310	460.6	1.49
TRNP - North	208	8,305.3	39.93	261	551.4	2.11
LNWA	506	424.0	0.84	149	157.1	1.05

* See also Appendix F in Addendum B to Exhibit 158.

Sulfur dioxide PSD baseline emission rates for most wells or flares were extrapolated from the five-month study period of the NDDH's Williston Basin Study (WBS). The NDDH improved the PSD baseline rates for the Little Knife Field, which is located 25 km southeast of TRNP-NU and generally east of TRNP-Elkhorn Ranch Unit. [77] Periods of available data representing normal oil production (normal operations) for calculating sulfur dioxide emission rates of wells in the Little Knife Field varied from three to ten months during 1977-78 due to Governor Link's March-1978 request to curtail daily oil production until the Little Knife Gas Plant began operations (which was July 1978). In a few instances, periods of available data were less than three months. (Note: Region 8 did not use the NDDH's improved rates for sources in the Little Knife Field. [78])

In sum, the NDDH's method of estimating sulfur dioxide emission rates for oil production flares and treaters did not average short-term peak emissions as stated by EPA Region 8.

[77] See Appendix F to Addendum B to Exhibit 158, pages 50 and 51.

[78] Exhibit 84, pages 21 and 22.

4.9 The NDDH encouraged oil field sour gas recovery rather than flaring.

As early as 1979, the NDDH had promoted and accommodated tying oil wells to gas plants to reduce and recover sour gas as waste and to reduce hydrogen sulfide and sulfur dioxide emissions. For example, the Little Knife Gas Plant proceeded with construction prior to receiving a Permit to Construct and was able to start operating five months after receiving a Permit to Operate.

First-in-time inventories of hydrogen sulfide and sulfur dioxide emitted by oil well operations located within 50 km of TRNP and LNWA began in 1982 and were completed in 1983. An attachment [79] to NDDH correspondence in April 1983 states the following:

> "Prior to 1977-1978, the start of rapid growth in oil exploration and development in North Dakota, there were few wells that were not connected to natural gas processing plants, and oil wells were considered minor sources of air pollution."
>
> "Air pollution from unconnected oil wells in the form of hydrogen sulfide and sulfur dioxide has increased and oil wells can now be considered a major [categorical] source of emissions. The reasons for this are several. The oil exploration and development boom which began in 1977-1978 reached its peak in 1981. Many of these wells were accompanied by sour natural gas in the range of 10 to 25 percent hydrogen sulfide. The tying of these new wells into natural gas processing plants has been delayed resulting in flaring of the gases and the increase in emissions of hydrogen sulfide and sulfur dioxide from oil wells. The delays were due to several factors, including the large number of new wells, time requirements for construction of new gas plants and pipelines, time requirements of the PSD permitting process for gas plants, and a surplus of gas and decline in demand for it."

Regarding its concerns over minor source emissions, EPA Region 8 continues by stating:

> "EPA is concerned that *direct use* of WBS 1987 to 1988 data will overestimate base year emissions and the amount of increment expansion credit. The concern can be seen by referring to ... the Statewide oil production data provided by the North Dakota Industrial Commission ... [T]he volume of oil produced in 1988 is nearly double that produced in 1976–77. Obviously, the much lower production rates in 1976–1977 would have tended to produce lower emissions than the State's estimate using the 1988 WBS data." [80] (italics added)

---

[79] See *North Dakota State Department of Health Strategy for Mitigating Effects of Oil and Gas Development near Class I Areas* dated 1983, attachment 8. Examples of ongoing tying of wells to sour gas processing plants are provided in *Individual Oil Wells in the Vicinity of Theodore Roosevelt National Park and Discussion of Wells Being Tied Into Gas Processing Plants*, attachment 7.

[80] See Exhibit 84, page 22.

Direct use of the WBS 1987 and 1988 oil-production emissions data may have overestimated PSD base year emissions. However, the NDDH did not apply a direct use of the WBS study; e.g., "only wells that actually existed and were producing during 1976-77 were included" in PSD base years' sulfur dioxide emissions inventory. [81]

Annual oil production in North Dakota is shown in figure 11. Oil production peaked in intervening years between PSD base years and current years (2000-01). Average annual production was 24.0 million barrels during PSD base years and 32.2 million barrels during current years, or PSD base production was 74.5% of current production.

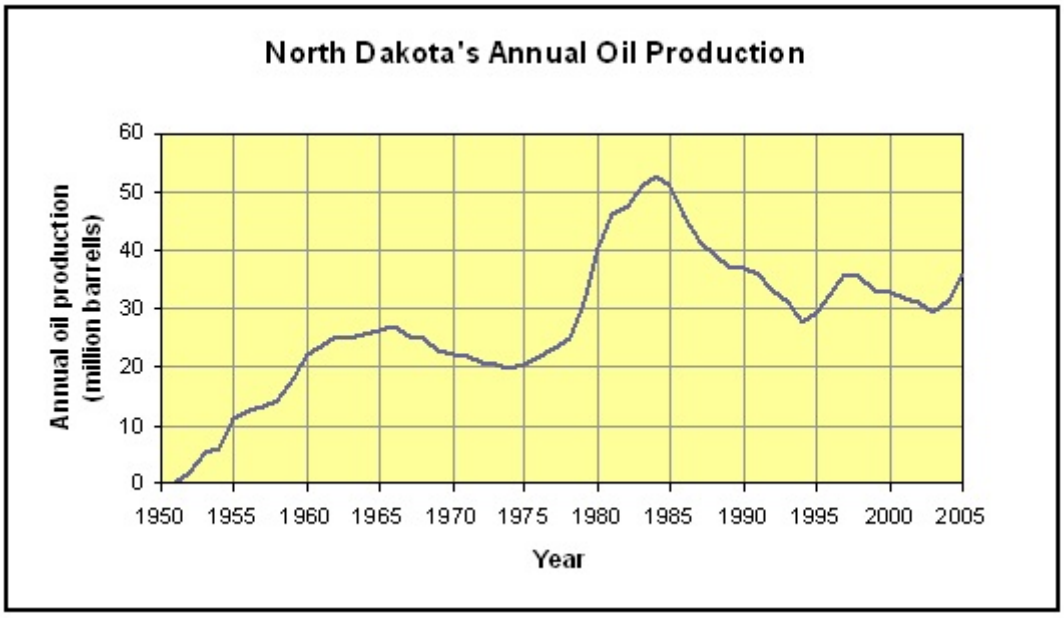


Figure 11. (Data taken from the web site for the Oil and Gas Division, North Dakota Industrial Commission.)

Annual average production per oil well is shown in figure 12. Annual oil production per well has not increased, but instead declined since 1978. So, implicitly, hydrogen sulfide or sulfur dioxide emissions for individual wells on average has also declined since 1978. Net total sulfur dioxide emissions of oil and gas flares and treaters have declined substantially since 1983. (See figure 2.)

A chronological list of sour natural gas and sulfur recovery plants is shown in table 4. (Owners or operators and names of some plants changed throughout the years.) The plants were constructed to reduce and recover sour gas rather than waste it by flaring and to reduce and recover sulfur in hydrogen sulfide and sulfur dioxide.

[81] See Appendix F to Addendum B to Exhibit 158, pages 48 and 49.

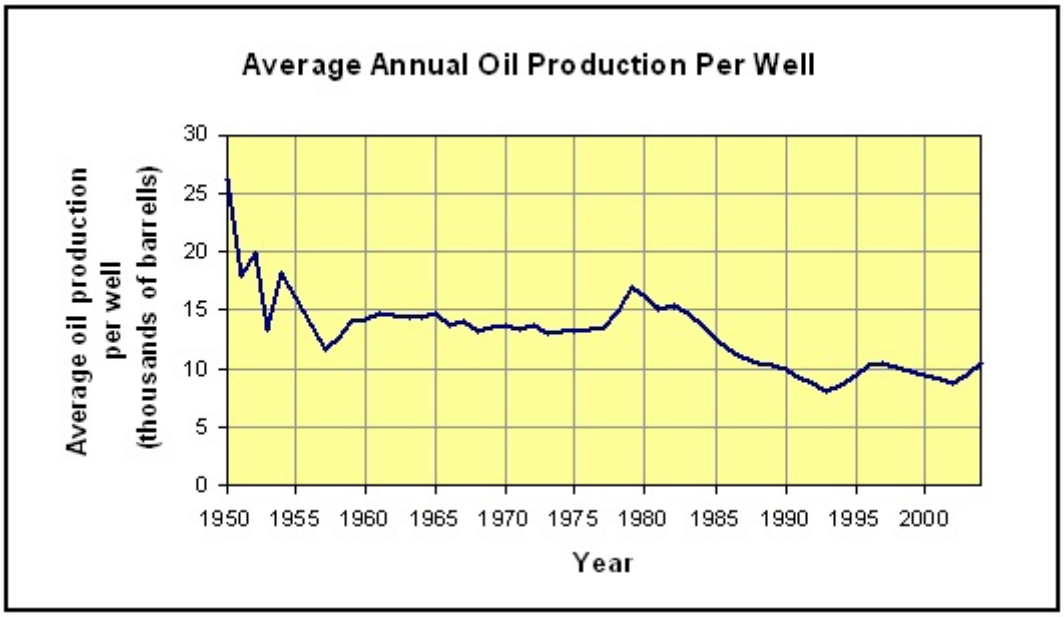


Figure 12. (Data taken from the web site for the Oil and Gas Division, North Dakota Industrial Commission.)

Table 4. Start-up Chronology of Sour Gas and Sulfur Recovery Plants

Gas and Sulfur Recovery Plants	Startup mo-yr	Operating
Tioga Gas	unk-54	yes
Lignite Gas	unk-61	yes
Boxcar Butte	unk-76	no
Little Knife Gas	Jul-78	yes
Shell-Oil Gas	Jul-79	no
Perry Petrolane	unk-unk	no
Killdeer Gas	unk-80	no
Grasslands Gas	Apr-80	yes
Teddy Roosevelt	Nov-80	no
Trenton Gas	unk-81	no
Whitetail Gas	never built	
Temple Gas	unk-85	no

See Appendix C to Addendum B, Exhibit 158.

4.10 In sum, EPA Region 8's larger PSD increment-affecting SO_2 emissions were due to several factors.

1st. (See sections 4.1 and 4.6 and table 2.) EPA Region 8's major-source current-period inventory included the bypass and the main stacks at the Great Plains Synfuels plant. But the bypass stack is only used when the main stack malfunctions, so including both stacks is a double counting. Otherwise, the sources included by both the NDDH and Region 8 in current-period and PSD baseline inventories were similar.

2nd. (See sections 4.1, 4.2 and 4.3.) Sulfur dioxide emission rates, as an average rate in tons per year reduced to pounds per operating hour and as 90th percentiles of 24-hour averages of hourly CEMS emissions data, are statistical compilations derived from the same EPA Acid Rain Program emissions data archive. Many of the state's major sources do not operate each and every hour throughout the year.

The NDDH used sulfur dioxide emission rates, as an average rate in tons per year (alternately lb/op-hr), for all major sources emitting sulfur dioxide, so as to provide "apples-to-apples," comparable rates between current-period and PSD baselines and among all sources in these inventories. These rates results in negligible risk for underestimating sulfur dioxide concentrations with models.

EPA Region 8 used 90th percentiles of 24-hour averages of hourly CEMS emissions data for electric utility boilers. And it also used an average of total annual emissions during all 365 days for all other major sources in both inventories. "Where CEMS data were not available to calculate a peak-to-mean ratio (e.g., the gas plants and the refinery), short term emissions were calculated by dividing the annual average emissions over 365 days." [82] But no CEMs data were available for the Neal Station and the Beulah Station; the sulfur dioxide 24-hour emission rates for these sources were also annual emissions divided by 365 days. [83] Region 8's emission rates are not "apples-to-apples" comparable rates among sources.

Neither the 90th percentile of 24-hour averages of hourly emissions nor the annual emissions during all 365 days are included in the implementing regulations definition for "actual emissions". The definition's command is either allowable emissions (per an enforceable permit) or total tons per year during hours of operation. The state's major sources do not ordinarily operate each and every hour throughout the year.

3rd. (See sections 4.1 and 4.4.) Following EPA's lead in preambles to rules (and also quoted by Region 8), the NDDH evaluated electric utility utilization of rated maximum boiler heat input as an indicator of demand for power. It concluded that, in some

---

[82] See Exhibit 84, pages 32 and 33.

[83] Id., page 38.

instances, demand for power prior to December 19, 1977 (the PSD minor source baseline date for western North Dakota) did not represent source normal operations.

Citing an EPA ***draft*** 1990 NSR workshop manual and EPA letters, Region 8 indicated that two years other than the two immediately preceding the PSD minor source baseline date can be used as representing normal operations only when retooling, major industrial accidents or labor strikes had occurred. [84] Accordingly, its draft modeling applied a more stringent application of PSD baseline source normal operations.

4th. (See sections 4.1 and 4.5.) For electric utility boilers, the NDDH calculated current-period sulfur dioxide emission factors as a sulfur balance between sulfur dioxide CEMS emissions data and corresponding coal consumption and coal sulfur-content data. The calculations revealed that the AP-42 sulfur dioxide emission factor of 30 is too low for most state utility lignite-fired boilers and does not reflect actual sulfur retention in boiler slag and ash. The NDDH used factors varying from 27.0 to 38.7 rather than 30, which resulted in larger sulfur dioxide PSD baseline emission rates for some boilers.

EPA Region 8 used its agency's AP-42 sulfur dioxide emission factor for lignite-fired boilers, although each lignite-fired boiler, in combination with coal quality, has an unique sulfur dioxide emission factor. [85] As a consequence, Region 8's current-period and PSD baseline sulfur dioxide emission rates for these sources distort and widen the difference between current-period and PSD baseline emissions.

5th. (See sections 4.1, 4.2 and 4.6.) Ninetieth percentile emission rates are larger than actual emissions as tons per year. [86] EPA Region 8's use of (a) 90th percentiles of 24-hour-averaged hourly CEMS sulfur dioxide emissions for sources constructed after PSD baseline and (b) annual daily averaged emissions for sources permitted to construct before the PSD major source baseline date distorts and widens the difference between current-period and PSD baseline emissions. Region 8's emission rates for current-period and PSD baseline sources are not "apples-to-apples" comparable rates.

6th. (See sections 4.7, 4.8 and 4.9.) Available data for oil production flares and treaters during PSD baseline, as well as late-1970s State policies on capturing waste sour gas, confound calculation of baseline sulfur dioxide emissions from these sources. The inexact baseline data and emission rates are moot issues because these source emissions did not significantly contribute toward consumption over or under the 24-hour class I increment.

---

[84] See Exhibit 57, pages 21 and 22.

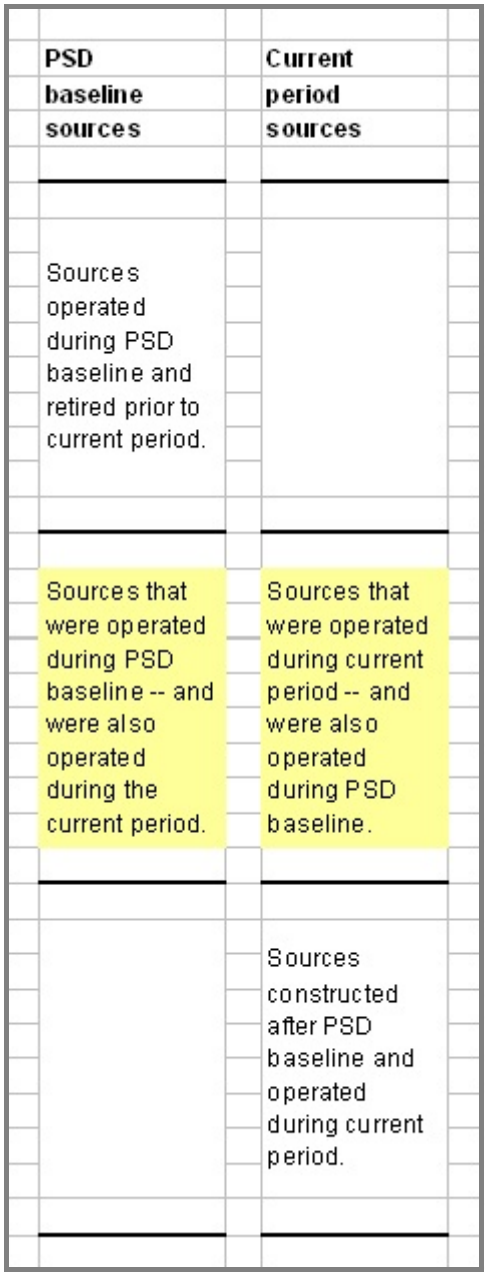
[85] See Exhibit 120. See also Exhibit 112, pages 21 (line 20) through 24 (line 7), and Exhibit 83, page 21 through 34.

[86] See Exhibit 33, table 3.

Part 5

Model Results
Uncertainty and Inaccuracy.

5.1 Meteorological data and actual concentrations are available for the current period, but not for PSD baseline.



When using models to estimate time trends in sulfur dioxide concentrations since the PSD baseline years, two inventories of sulfur dioxide emissions are required – current and baseline.

"The accuracy of model estimates varies with the model used, the type of application, and site specific characteristics." [87]

"In all applications of models an effort is encouraged to identify the reliability of the model estimates for that particular area and to determine the magnitude and sources of error associated with the use of the model. The analyst is responsible for recognizing and quantifying limitations in the accuracy, precision and sensitivity of the procedure. ... Both space/time pairing of estimates and measurements and unpaired comparisons are recommended." [88]

"Basically, ... (2) the models are reasonably reliable in estimating the magnitude of the highest concentrations occurring sometime, somewhere within an area. ... However, estimates of concentrations that occur at a specific time and site, are poorly correlated with actually observed concentrations and are much less reliable." [89]

Uncertainty and inaccuracy in model-estimated concentrations are due to several factors. Uncertainty and inaccuracy, for example, are due to (1) inadequate spatial and temporal resolution of input meteorological and other data and (2) imperfect algorithms for: (a) actual atmospheric

[87] See Exhibit 132, section 10.1.3(a). This exhibit is EPA's *Guideline on Air Quality Models*, which is Appendix W to 40 CFR Part 51.

[88] Id., section 10.1.3(b).

[89] Id., section 10.1.2(a).

air in motion (pollutant trajectories), (b) actual mixing depth, and (c) actual physical processes (vertical and horizontal pollutant dispersion and pollutant depletion via deposition and chemical transformation).

When assessing the accuracy of model-estimated concentrations, "known" results of sulfur dioxide actually emitted and then transported and dispersed by actual occurring meteorology are needed. The "known" results are typically measured concentrations actually recorded. [90] (See also sections 3.6 and 3.7.)

Prior to 1980, no reliable actual concentrations of sulfur dioxide are available for locations in the NU and SU of TRNP or for other locations in the state. The year 1980 was the first year of reliable actual concentration data throughout the entire year for the two units of the park. The first year of reliable actual concentration data throughout the year for the LNWA is 2004.

However, two models are necessary and were used by the NDDH and EPA Region 8 for estimating sulfur dioxide concentrations. Calmet uses algorithms to spatially and temporally interpolate input meteorological data and prepares output wind fields and other data that Calpuff uses as input, along with sulfur dioxide emissions and other data. Thus, Calmet and Calpuff are executed in tandem.

So when assessing accuracy of model-estimated concentrations, mesoscale meteorological data and other data are needed for the period of emitted sulfur dioxide. Mesoscale meteorological data and other data suitable for Calmet input are available for the period of the current-emissions inventory, but not for the period of the PSD baseline emissions inventory.

In sum, only current-period mesoscale meteorological data and actual sulfur dioxide concentrations are available for accuracy testing of model-estimated wind fields and model-estimated concentrations.

A three point model performance accuracy was completed: (1)(a) NOAA/NWS Rapid Update Cycle (RUC) data and (b) MM5 data were compared to NWS surface and upper air wind observations within the NDDH's Calmet modeling domain (section 5.2), [91] (2) using RUC and MM5 data as input, Calmet output wind fields were compared to independent wind-energy meteorological tower wind data within the NDDH's modeling domain (section 5.2), and (3) Calmet–Calpuff-estimated sulfur dioxide concentrations were compared to actual sulfur dioxide concentrations (sections 5.3 – 5.8).

---

[90] See *Performance Measures and Standards for Air Quality Simulation Models*, EPA-450/4-79-032, October 1979, Office of Air Quality Planning and Standards, Research Triangle Park, NC, pages VI-2 and VI-3.

[91] See PowerPoint presentation by Bret Anderson, EPA Region 7, dated May 18, 2005, and quoting section 9.3 of EPA's *Guideline on Air Quality Models* (Appendix W to 40 CFR Part 51 or Exhibit 132) on page 7, at http://cleanairinfo.com/modelingworkshop/presentations/MM5_Anderson.pdf

5.2 Model performance anchors to spatial and temporal accuracy of the meteorological input data.

The Calpuff model requires three-dimensional wind field data, other weather data, source pollutant emissions data, source stack data and additional miscellaneous data. The three-dimensional wind field data are prepared prior to Calpuff execution using the Calmet diagnostic meteorological model.

Calpuff's estimated sulfur dioxide concentrations reflect the accuracy of raw meteorological data, such as wind speed and direction, the realism of Calmet's computational algorithms, and the accuracy of Calmet's computed wind fields.

NWS observations and MM data sets were used by the NDDH and EPA Region 8 in their respective 2002 and 2003 protocols. NWS observations and RUC data sets prepared by WindLogics, Inc., (see section 8.5) were used by the NDDH in its MOU Protocol. Accuracy analyses of (1) the RUC and MM5 data as input to Calmet and (2) the output of Calmet were completed.

5.2.1 Accuracy of RUC Data as Input to and Output from Calmet

WLI, for example, completed a statistical comparison of RUC surface winds with NWS surface winds; the results indicate good agreement. [92] "[T]he mean absolute error is less than 0.6 m/s as compared to a typical forecast [MM5] error of 2 m/s or more." [93] And in response to a NDDH 2002 review of RUC data implicitly cited by EPA, [94] ENSR Corporation completed a comparison of RUC upper air data with NWS Bismarck upper air data; these results also indicate good agreement. [95]

The NWS measures surface winds each hour and upper air winds every twelfth hour. The average distance between NWS surface stations in the modeling domain is about 135 kilometers, and the average distance between NWS upper air (rawindsonde) stations is more

---

[92] See Addendum D to Exhibit 158, which is titled *A Comparison of NOAA RUC Analysis Surface Winds and ADAS-Enhanced RUC Analysis Winds with Surface Observations*.

[93] Id., page 9. See also page 4.

[94] "A comparison of the wind speeds used in the consultants' modeling with those that were actually observed at the time showed that the modeled wind speeds were consistently higher than measured values. This overestimate of wind speeds was particularly apparent during stagnant conditions when predicted concentrations would be expected to occur." See Exhibit 84. Page 9.

[95] See section 3 of Exhibit 95, Vol. 3, Tab B, which is titled *Revised CALPUFF Analysis with Year 2000 MM5 Meteorological Data ...* by ENSR Corporation for Basin Electric Power Cooperative.

than 400 km. [96] NWS surface and upper air wind observations are often not representative of spatially varying winds over great distances. (See section 5.9. [97]) Interpolation of NWS observations having "large spatial and temporal gaps [by Calmet to the Calmet grid of 10 km or less] can lead to artificial reductions of scalar wind speeds by averaging vector winds when the meteorology is changing in time and space." [98]

The accuracy of Calmet output wind speeds when using (1) NDDH's NWS-only observations (without MM5 or RUC) with its 2002 draft protocol and (2) WLI's RUC data with ENSR's protocol has been examined. [99] Some user inputs in the NDDH and ENSR protocols differed, such as values for Calmet variables R1 and R2.

Calmet vertical-layered output wind speeds at grid points nearest five meteorological towers were compared to wind speeds at the towers. Days were divided into eight 3-hour blocks and an average speed calculated for each 3-hour block to smooth the stochastic hour-to-hour magnitude of wind speeds. [100] An example comparing Calmet wind speeds to Killdeer-tower wind speeds is shown in figure 13. For all hours of the day, Calmet output wind speeds are lower than observed meteorological-tower wind speeds, whether using (1) NWS-only observations or (2) RUC data. However, the Calmet RUC-derived wind speeds are systematically higher than the Calmet NWS-derived wind speeds.

In addition, Calmet output wind speeds that are less than 5 m/s, whether using NWS-only observations or RUC data, occur more frequently (larger percentage of all wind speeds) than tower wind speeds less than 5 m/s. [101] An example is shown in figure 14.

---

[96] See map in Addendum B to Exhibit 158, page 9.

[97] Surface and 900-mb animations of raw RUC 40-km wind data have been created; for surface animations, see http://ndhealth.gov/AQ/Dockets/PSD/Clip%201.mov and for 900-mb animations, see http://ndhealth.gov/AQ/Dockets/PSD/Clip%203.mov.

[98] See Exhibit 95, Vol. 3, Tab E, which is titled *Comparison of CALMET Wind Speed Predictions with Measurements from Wind Energy Meteorological Towers in Western North Dakota* by ENSR Corporation for Basin Electric Power Cooperative, page 1-1. See also Attachment 6 to Tab B, Volume 3, Exhibit 95. "It is clear that the time-interpolated values [of NWS observed wind speeds] fail to capture the details of the weather front as it moves through the state. Neither the directions nor the speeds of the time-interpolated wind values agree with the results of the NOAA [RUC] analysis, which is the best fine-scale data available from the National Weather Service. Time-interpolated values should not be used when better data, [sic] more physically consistent data sources such as the NOAA [RUC] analyses, are available." (Id., page 3.)

[99] Id.

[100] Id., figures 4-1 through 4-5.

[101] Id., figures 4-6 through 4-10.

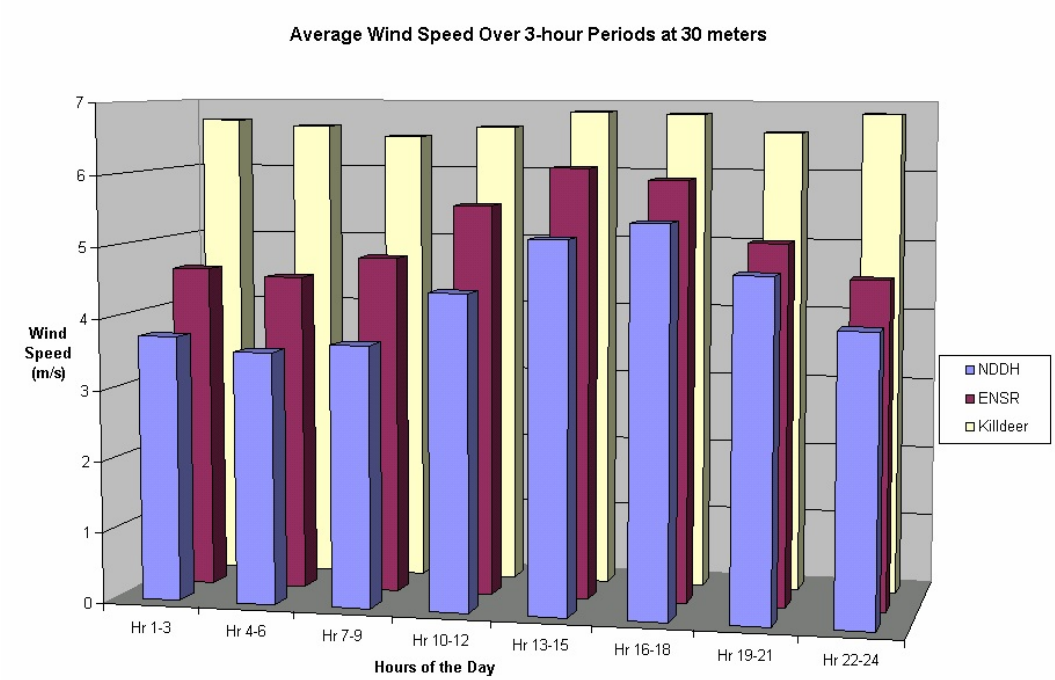


Figure 13. Near Killdeer (January – December 2001, figure 4-4 in Vol. 3, Tab E, Exhibit 95 by ENSR Corporation)

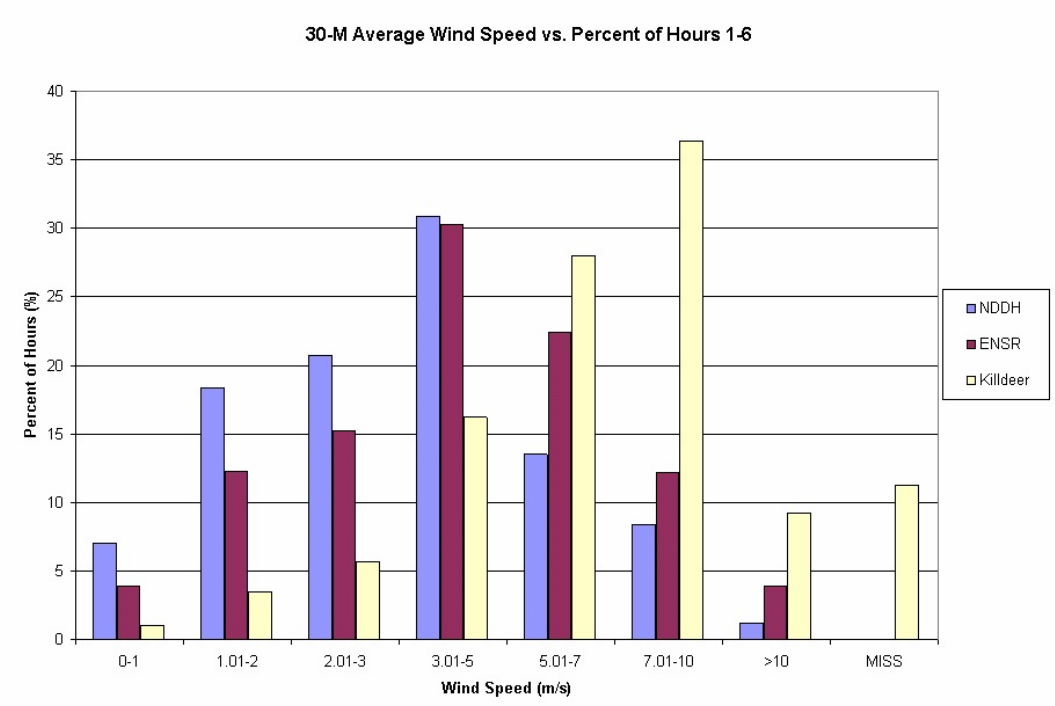


Figure 14. Near Killdeer (January – December 2001, figure 4-9 in Vol. 3, Tab E, Exhibit 95 by ENSR Corporation)

In sum, the graphed data in figure 13 illustrate that use of (1) NWS observations alone and (2) RUC data with Calmet underestimates wind speeds and, therefore, implicitly illustrate that use of these NWS and RUC data overestimates actual sulfur dioxide concentrations, since plume dilution is proportional to wind speed. The graphed data in figure 14 illustrate a bias in Calmet output wind fields, since (a) Calmet output wind speeds less than 5 m/s occur more frequently than tower wind speeds less than 5 m/s and, correspondingly, (b) Calmet output wind speeds greater than 5 m/s occur less frequently than tower wind speeds greater than 5 m/s.

5.2.2 Accuracy of MM5 Data as Input to and Output from Calmet

The only known comparison of MM5 data with NWS surface wind data in the NDDH's modeling domain is provided in the report for EPA Region 8 by Alpine Geophysics. [102] These MM5 data, for year 1994, were used by Region 8 in its draft modeling reported in 2003. [103] Alpine Geophysics used a performance metric termed "Index of Agreement" as an indicator of agreement between MM5 forecast surface winds and NWS observed surface winds. During the year, the index has a large variability ranging from about 0.15 to 0.95 with an average of 0.71. An index value of 1 suggests prefect agreement; the index has no physical units and is not directly comparable to the mean absolute error, which as units of m/s.

The accuracy of Calmet output wind speeds per EPA Region 8's draft protocol, which used Alpine Geophysics 1994 MM5 data, has also been examined. [104] Calmet three-dimensional output wind fields were obtained from Region 8 and compared to wind data from two meteorological towers in western North Dakota. Calmet vertical-layered output wind data for the Calmet grid point nearest the tower were used, and data from layers at 36 m and 75 m were averaged to represent an effective height of about 55 meters so as to match the height of a wind monitor on the tower.

Days were divided into eight 3-hour blocks and an average speed calculated for each 3-hour block. [105] An example comparing Calmet wind speeds to Wilton-tower wind speeds is shown in figure 15. For all hours of the day, Calmet output wind speeds are substantially lower than observed Wilton-tower wind speeds. In addition, Calmet output wind speeds that are less than 7 m/s occur more frequently (larger percentage of all wind speeds) than Wilton-tower wind speeds less than 7 m/s. [106] An example is shown in figure 16.

---

[102] See Exhibit 135 titled *Annual Application of MM5 to Support 1994 Calpuff Air Quality Modeling* by Alpine Geophysics for EPA Region 8.

[103] See Exhibit 84, pages 4 – 7.

[104] See Attachment 5 to Tab E in Vol. 3 for Exhibit 95, which is titled *Comparison of Wind Tower Data With EPA's MM5 CALMET for October – December 1994.*

[105] Id., figures 3 and 4.

[106] Id., figures 5 through 12.

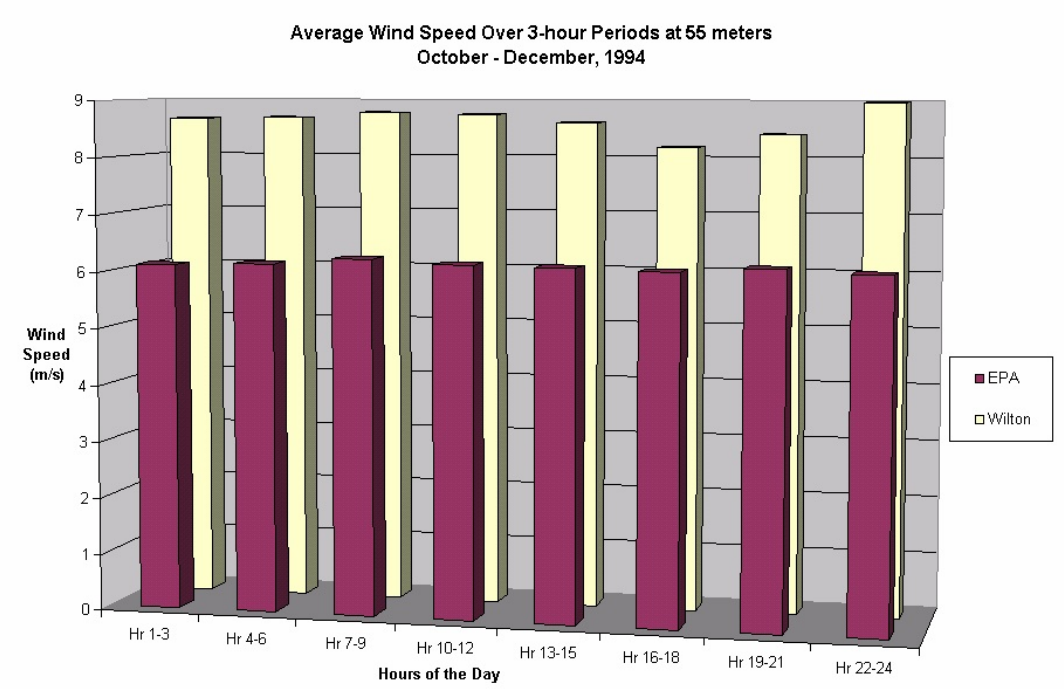


Figure 15. Near Wilton (Figure 3 in Attachment 5 to Vol. 3, Tab E, Exhibit 95 by ENSR Corporation)

In sum, the graphed data in figure 15 illustrate that use of MM5 data with Calmet also underestimates wind speeds and, therefore, implicitly illustrate that use of these MM5 data overestimates actual sulfur dioxide concentrations, since plume dilution is proportional to wind speed. The graphed data in figure 16 illustrate a bias in Calmet output wind fields, since (a) Calmet output wind speeds less than 7 m/s occur more frequently than tower wind speeds less than 7 m/s and, correspondingly, (b) Calmet output wind speeds greater than 7 m/s occur less frequently than tower wind speeds greater than 7 m/s.

5.2.3 Comparison of Calmet Output using RUC or MM5

Only two of the five meteorological towers were operated at times when collected data could be used to compare to Calmet output whether using RUC or MM5 data. As shown with Wilton-tower wind data in figure 17, Calmet output wind speeds that are less than 5 m/s, whether using NWS-only observations or RUC data, occur more frequently (larger percentage of all wind speeds) than tower wind speeds less than 5 m/s.

The data graphed in figures 13, 16 and 17 clearly suggest that (1) spatially- and temporally-deficient NWS observations alone (without RUC or MM5 data) should not be used as input to Calmet and (2) NWS observations should not be used to substantially replace RUC or MM5 data in the Calmet stepwise calculation of three-dimensional wind fields.

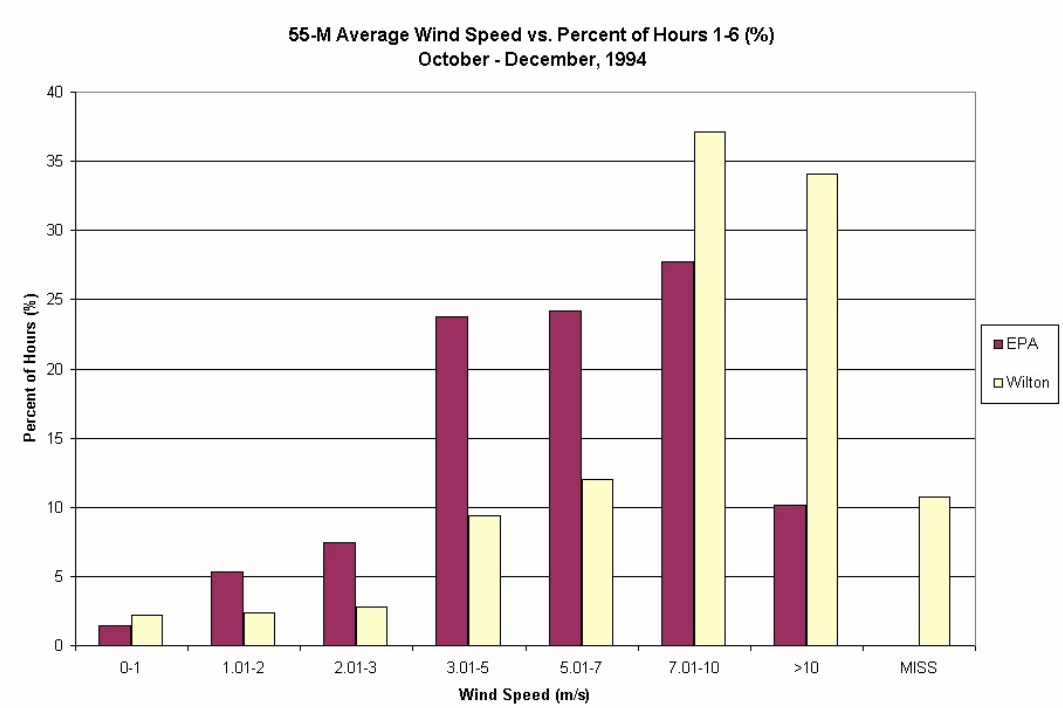


Figure 16. Near Wilton (Figure 5 in Attachment 5 to Vol. 3, Tab E, Exhibit 95 by ENSR Corporation)

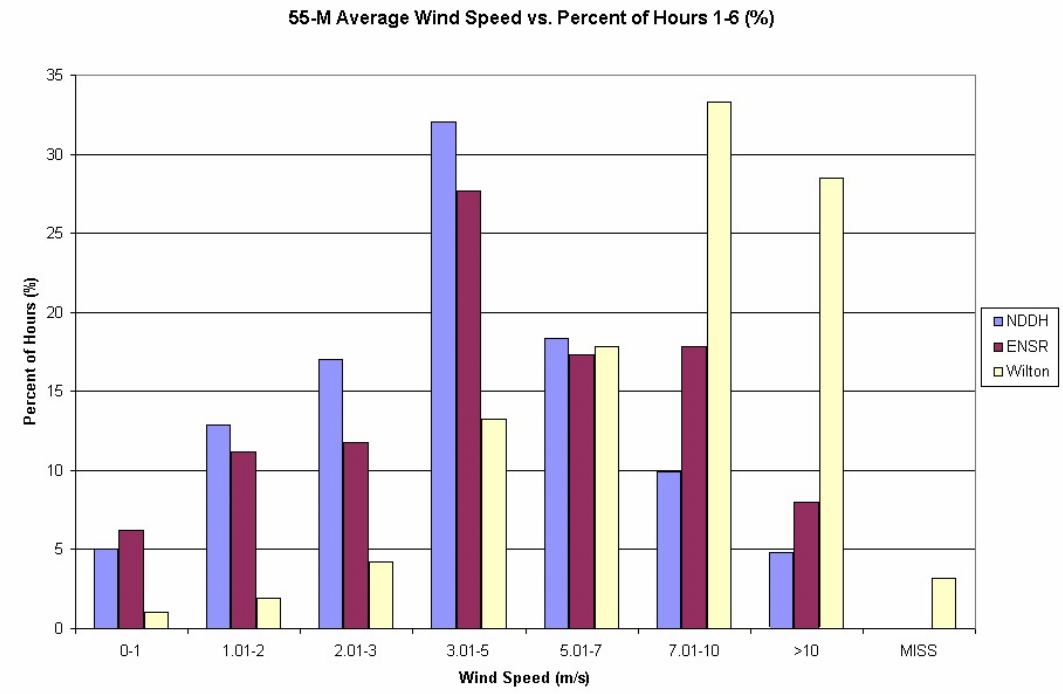


Figure 17. Near Wilton (Figure 4-7 in Vol. 3, Tab E, Exhibit 95 by ENSR Corporation)

5.3 RUC data as Calmet input result in higher estimated SO_2 concentrations.

The State's MOU Protocol was executed using three year-2002 meteorological data sets – RUC mesoscale meteorology with NWS surface and upper air data, MM5 with the NWS data [107] and the NWS data without RUC or MM5 data. All other Calmet data inputs were identical when using RUC and MM5 data; however, when using only NWS observations, R1 was increased to 46 km, R2 to 91 km, RMAX1 to 200 km and RMAX2 to 800 km. All Calpuff data inputs were also identical.

Daily model-estimated sulfur dioxide concentrations [108] for a model receptor at the site of the monitor in the NU of TRNP were rank ordered from highest to lowest. Likewise, daily averaged actual concentrations were rank ordered. The highest of the 365 model-estimated concentrations was paired with the highest of the 365 actual concentrations, etc. Figure 18 shows differences (estimated minus actual in ug/m3) for the first pair through the 25th pair.

These performance accuracy tests illustrate use of RUC data results in estimated concentrations larger than use of MM5 or NWS observations. The tests also illustrate that the models are overestimating actual concentrations, even with the background concentration set to 0 ug/m3. The estimated background concentration for sulfur dioxide is 1.5 ug/m3. (See section 8.4.)

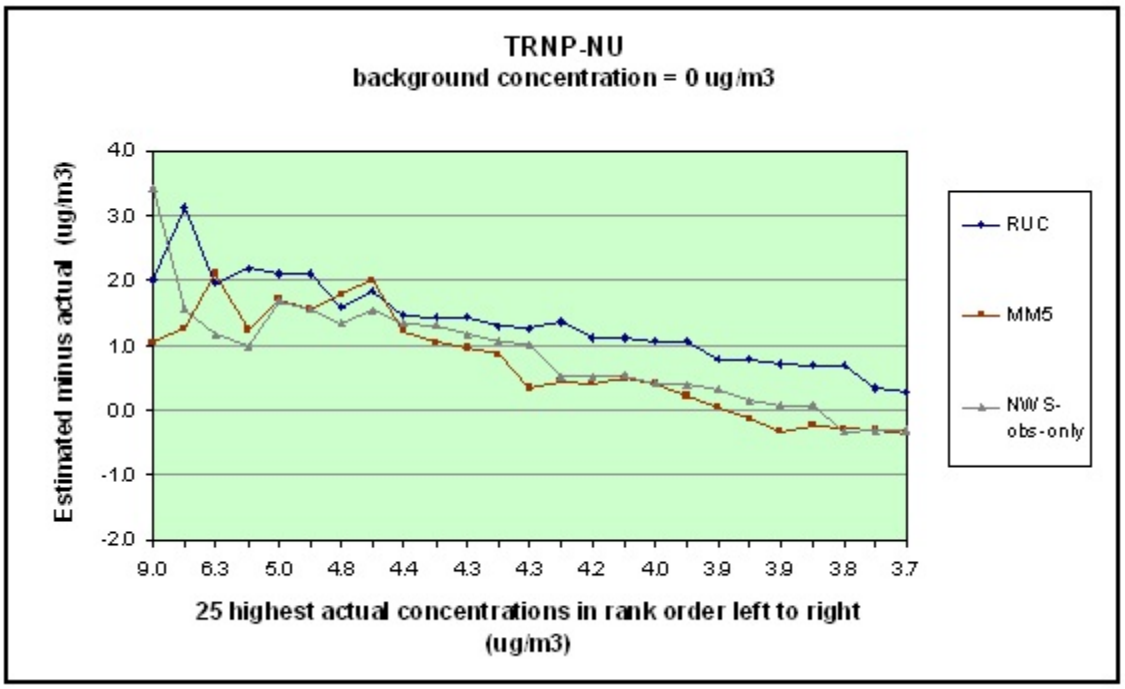


Figure 18. Differences in Matched SO2 24-hr Modeled and Monitored Concentrations

[107] MM5 data were provided by Mr. Tim Allen, FWS, in June 2006 without documentation.

[108] Calmet (version 5.53a) and Calpuff (version 5.711a) models, as approved by EPA via Appendix W to 40 CFR Part 51, were used for these performance accuracy tests.

5.4 Model-estimated SO_2 concentrations using MM5 blended with NWS observations are not better than RUC with no blending of NWS observations. (See discretionary options 5 and 6 per MOU.)

Wind speed and other meteorological factors govern levels of actual concentrations measured with monitors. (See figures 8 and A2.) In its draft May 2003 report, EPA Region 8 states:

> "From the [ENSR] Calpuff/Calmet input files it appears that [RUC data formatted as] MM5 data was used to *replace all* of the actual measured surface and upper air weather observations. This practice is not recommended by either IWAQM or the Calpuff model developer. MM5 as currently configured tends to overestimate wind speeds, particularly in the lower levels of the atmosphere, unless restrained by actual observations." [109] (italics added)

Small values for Calmet input variables R1 and R2 substantially replace NWS-measured surface and upper air wind observations with mesoscale meteorological wind data such as RUC or MM5. [110]

EPA Region 8 used MM5 wind data and NDDH's Calmet-input-ready NWS wind observations for its 2002 and 2003 draft modeling. [111] Region 8's input values for Calmet-variables R1 and R2 were 1 km and 10 km, respectively; [112] so the MM5 data substantially replaced the actual surface and upper air wind observations in contradiction to Region 8's comment quoted above.

The NDDH also used MM5 wind data and NWS wind observations with its 2003 modeling. Its input values for Calmet's R1 and R2 were 40 km and 60 km, respectively; so MM5 data were blended into NWS observations (as Calmet initial guess winds; i.e., variable IPROG = 14) consistent with the practice preferred by Region 8.

The State's MOU Protocol used RUC data in lieu of MM5 data. (See section 5.2.) The input values for Calmet-variables R1 and R2 are the same, or 10 km.

---

[109] See Exhibit 84, page 9.

[110] See Attachment 7 under Tab B, Vol. 3, Exhibit 95, which is an email message from EPA's OAQPS to Region 8 that indicates OAQPS preferences for certain Calmet input options. Region 8 had initially proposed values for R1 and R2 at 125 km and 100 km, respectively. But OAQPS recommended values for R1 and R2 as 1 km and 10 km, respectively, indicating that the large values loose "all the possible value-added [when using MM data] with Stage 1 processing, and are left with either a) a 1/r2 interpolated wind field ... or b) a 1/r2 wind field, with winds looking at any one level, identical to one another in the near-vicinity of observation locations."

[111] See, respectively, Exhibit 81, pages 10, 31 and 32 , and Exhibit 84, page 10.

[112] See Exhibit 84, page 12.

The Calpuff-in-tandem-with-Calmet performance test, as described in section 5.3 and shown in figure 18, was expanded by remodeling the MM5 mesoscale data and changing input values for Calmet variables R1 and R2 to 46 km and 91 km, respectively. In this manner, MM5 data do not *replace* NWS observations, but rather are blended into the NWS observations. All other Calmet and Calpuff input data were the same for this performance test. Results of the test are shown in figures 19, 20 and 21 as "MM5 rev R1, R2"; model-estimated concentrations were not adjusted by adding a background concentration.

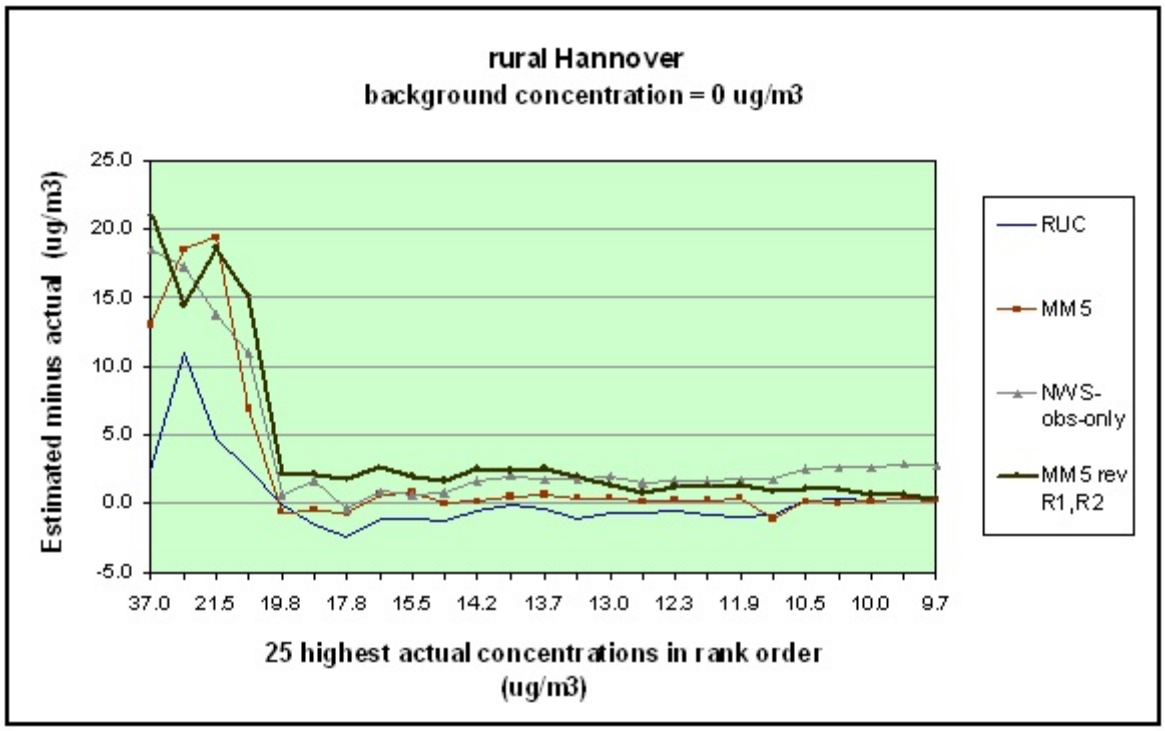


Figure 19. Differences in Matched SO2 24-hr Modeled and Monitored Concentrations

The estimated background concentration for sulfur dioxide is 1.5 ug/m3. When added to model-estimated concentrations, all data curves in figures 19, 20 and 21 (including figure 18) shift vertically by 1.5 ug/m3. The rural Hannover monitoring site is closer to most major sources of sulfur dioxide. (See figures 1 and 5.)

In sum, blending of the MM5 wind data into NWS wind observations with Calmet increased the bias in Calpuff estimated sulfur dioxide concentrations. Model-user-selected values for Calmet variables R1 and R2 should correspond to the quality of the meteorological data input to Calmet. When using the RUC mesoscale data, increasing values for R1 and R2 would distort Calmet output wind field patterns and result in reduced accuracy of the output winds. (See section 5.2.)

Finally, the State MOU Protocol's use of RUC data, and other model inputs, results in larger model-estimated concentrations in the NU and SU of TRNP than use of MM5 data.

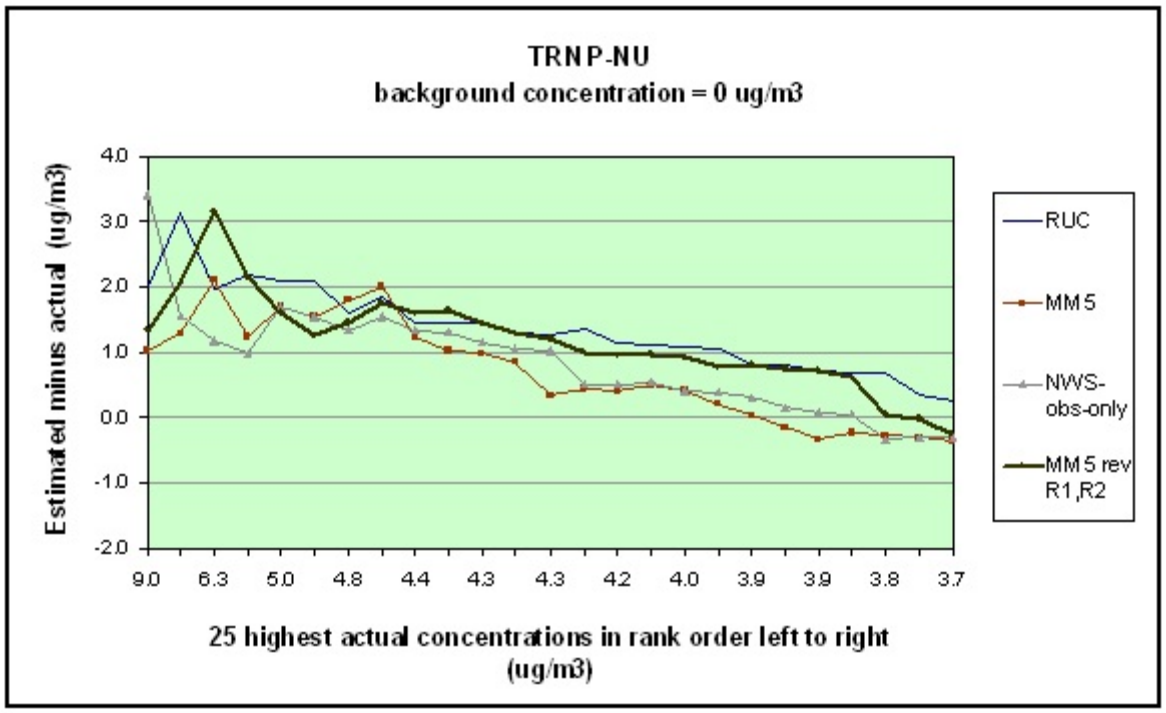


Figure 20. Differences in Matched SO2 24-hr Modeled and Monitored Concentrations

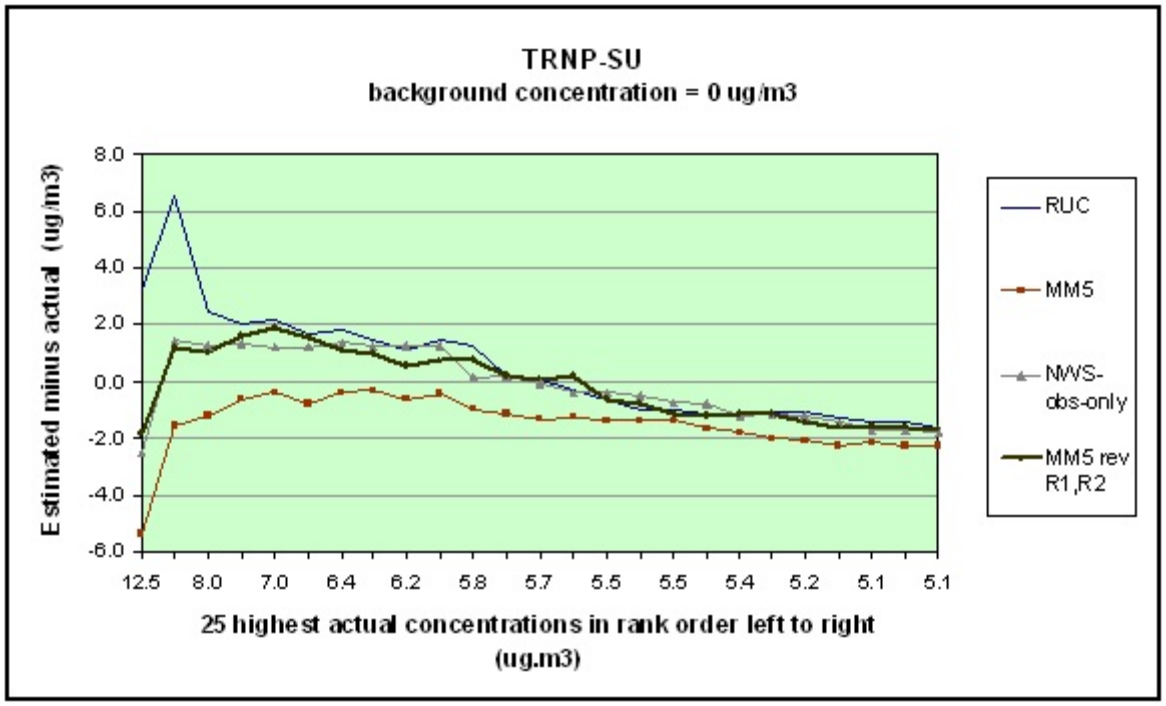


Figure 21. Differences in Matched SO2 24-hr Modeled and Monitored Concentrations

5.5 Normalized bias in largest model-estimated SO_2 concentrations is significant.

EPA has stated that

> "... models are reasonably reliable in estimating the magnitude of highest concentration occurring sometime, somewhere within an area ... However, estimates of concentrations that occur at a specific time and site are poorly correlated with actual observed concentrations ..." [113]

Mean bias (MB) is expressed as the mean difference between model-estimated concentrations matched paired-in-time with actual concentrations. Mean normalized bias (MNB) is expressed as the mean (e.g., i = 1,25) of normalized differences as follows.

$$MNB = \overline{[(C_M + bg) - C_A] / C_A}, \quad \text{where bg = background concentration}$$

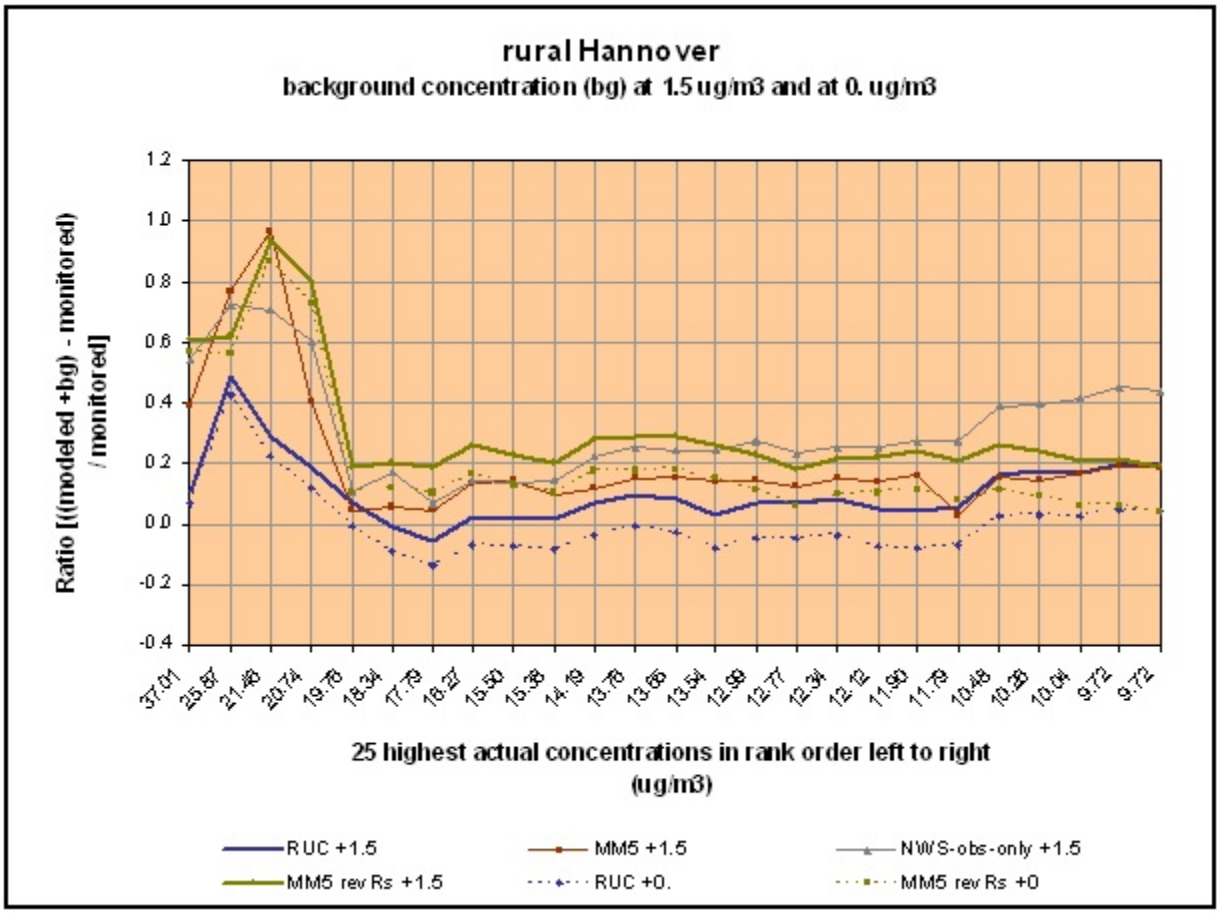


Figure 22. Normalized Bias Using Rank-order Matched Pairs of SO2 24-hour Concentrations

[113] Paragraph b, section 10.1.2 of Exhibit 132, which is Appendix W to 40 CFR Part 51.

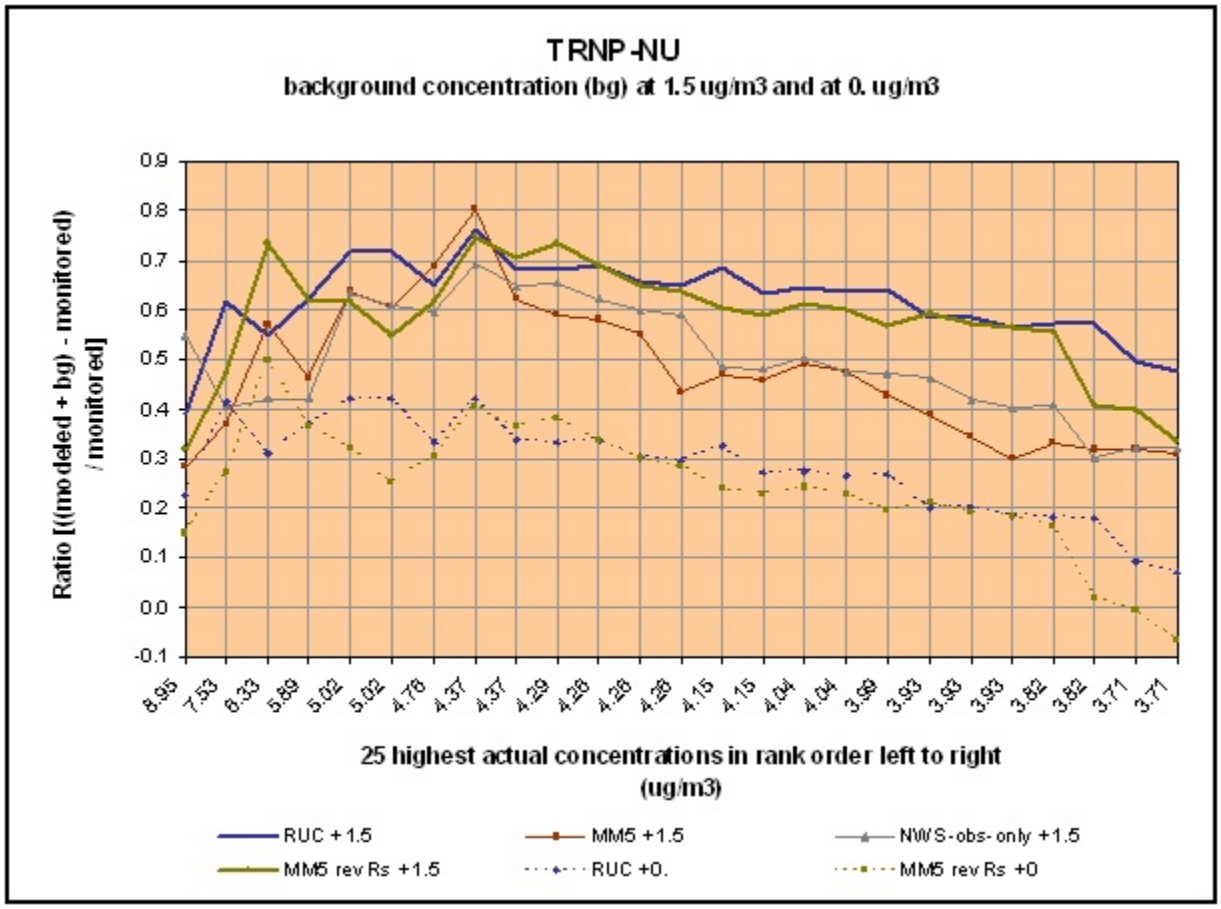


Figure 23. Normalized Bias Using Rank-order Matched Pairs of SO2 24-hour Concentrations

(See section 5.3.) Normalized bias of each of the 25 highest model-estimated 24-hour averaged sulfur dioxide concentrations rank-order matched (not paired in time) with respective 25 highest actual 24-hour averaged concentrations are shown in figures 22, 23, A5 and A6. Positive normalized bias (+) indicates a background-adjusted model-estimated concentration is larger than a paired actual concentration, and negative normalized bias (–) indicates a background-adjusted model-estimated concentration is less than a paired actual concentration.

In sum, the largest background-adjusted model-estimated sulfur dioxide 24-hour concentrations are greater than the largest actual sulfur dioxide 24-hour concentrations. The TRNP-NU monitoring site is located downwind (westward) of major sources and the rural Hannover site in the sulfur dioxide actual transport trajectory. (See figures 1 and 5.) The graphs of normalized bias of model-estimated concentrations at the rural Hannover and TRNP-NU monitoring sites are diametrical and illustrate larger normalized bias at the TRNP-NU site.

Finally, the largest background-adjusted model-estimated concentrations determined per the State MOU Protocol's RUC data, and other model inputs, are not less than the largest actual concentrations in the NU and SU of TRNP.

5.6 Model-estimated SO_2 concentrations can be bias corrected.

(See sections 5.3 and 5.5.) Since Calmet–Calpuff models are overestimating the largest current actual concentrations, whether using the State's MOU Protocol inputs or using alternate meteorological inputs, this question occurs.

Can model-estimated increases (deterioration) be bias corrected?

An illustration of mean normalized bias is shown in table 5. The illustration relates to plotted data in figures 22 and 23. Mean normalized bias approach 0.0 when differences between background-corrected model-estimated and actual concentrations approach 0.0. (Normalized bias approaches 0.0 when the ratio (R) of a background-corrected model-estimated concentration and an actual concentration ($(C_M + bg) \div C_A$) approaches 1.0. [114] The math expression for normalized bias is equivalent to the math expression R – 1 in the State's MOU Protocol. [115] When normalized bias is larger than 1.0, R is larger than 2.0.)

Data for mean normalized bias in table 5 indicate that:

✗ When using a background concentration of 1.5 ug/m3, mean normalized bias in model-estimated sulfur dioxide 24-hour concentrations is larger than 0.0 at each of the three monitoring sites, irrespective of meteorological input data.

  ✓ Consequently, the Calmet–Calpuff models are overestimating, on average, PSD increment consumption. [116]

✗ At the site of the rural Hannover monitor, the RUC mesoscale data result in substantially better agreement between model-estimated sulfur dioxide 24-hour concentrations and actual 24-hour concentrations. This monitoring site is frequently located in the actual trajectory path between the larger sulfur dioxide emissions of coal-fired utility boilers and occasionally located in actual westward trajectories that cross the TRNP PSD class I areas. [117]

✗ For all four types of meteorology data used, model-estimated sulfur dioxide 24-hour concentrations are less inaccurate at the site of the monitor in the SU of TRNP compared to the site of the monitor in the NU of TRNP.

---

[114] See sections 5.0 and 6.0 in Addendum C to Exhibit 158.

[115] See Appendix B to Addendum B to Exhibit 158.

[116] This assumes that bias in model estimates of PSD baseline concentrations is the equivalent to bias in model estimates of current concentrations. See Appendix B to Addendum B to Exhibit 158.

[117] See Appendix B to Addendum C to Exhibit 158. Some episodes of actual concentrations at rural Hannover, rural Dunn Center and at the NU and SU of TRNP occur concurrently.

Table 5. Mean Normalized Bias in Model-estimated SO2 24-hour Concentrations Unpaired in Time with Actual Concentrations *

	RUC	MM5	NWS obs-only	MM5 rev Rs
rural Hannover				
Highest 5	**0.23**	0.52	0.54	0.63
Highest 25	**0.11**	0.21	0.32	0.31
rural Dunn Center				
Highest 5	0.80	0.61	0.59	0.63
Highest 25	0.56	0.53	0.61	0.62
TRNP-NU				
Highest 5	0.58	0.46	0.49	0.55
Highest 25	0.62	0.47	0.50	0.58
TRNP-SU				
Highest 5	0.57	0.00 (0.13) **	0.28 (0.31)	0.31 (0.32)
Highest 25	0.29	0.01 (0.10)	0.21 (0.23)	0.21 (0.22)

* SO2 background concentration for model-estimated SO2 concentrations is 1.5 ug/m3.
** Numbers in parentheses are mean normalized error when different than MNB.

In sum, the question's answer is yes, or model-estimated concentrations can be bias corrected. The bias data in table 5 indicate that models and inputs are overestimating, on average, actual concentrations in the range of 0.00 to 0.80, or 0% to 80%. If a bias correction had been applied to model-estimated deterioration in section 8 of the NDDH's MOU Protocol Results Report (Addendum C to Exhibit 158), the model-estimated deterioration after PSD baseline would have been less than indicated.

This model accuracy analysis, as well as analyses in sections 5.3 and 5.4, used NDDH's "actual emissions" as an average of annual tons (tons per year) during source annual operating hours. (See table 2.) EPA Region 8's current-period (2000-01) sulfur dioxide emission rates are 22.8% larger than the NDDH's current rates. **If EPA Region 8 had completed a similar bias analysis, using the same actual sulfur dioxide concentrations, and applied a bias correction to its model-estimated deterioration, its bias-corrected, model-estimated deterioration would likely have converged toward the NDDH's estimated deterioration.**

5.7 The Calmet–Calpuff models are poor at matching day-to-day actual SO_2 concentrations.

(See section 5.3.) Three meteorological data sets were used with Calmet: RUC mesoscale meteorology with NWS surface and upper air data, MM5 with the NWS data, and the NWS data without RUC or MM5 data. Days of highest model-estimated concentrations and highest observed concentrations are listed in tables 6a and 6b. For example, the largest actual concentration (rank = 1) during 2002 at the monitoring site in the SU of TRNP occurred on Julian Day 248.

In sum, the days of highest 24-hour estimated concentrations using the models do not correspond in rank order with the days of highest monitored concentrations.

Table 6a. Julian Days (2002) of 25 Highest Sulfur Dioxide Concentrations

TRNP-North Unit

Julian Day 1 = January 1, Julian Day 365 = December 31

Rank		Model-estimated using:		
1=highest, etc.	Actual	RUC	MM5	NWS Obs *
1	73	73	73	**219**
2	66	66	66	**100**
3	83	**199**	**219**	238
4	80	50	**199**	153
5	39	**75**	**75**	**336**
6	95	83	337	39
7	78	39	139	174
8	29	116	**100**	159
9	92	294	**301**	294
10	302	337	83	230
11	69	137	116	**75**
12	293	**219**	153	250
13	337	250	39	216
14	79	74	127	300
15	116	293	189	127
16	153	**301**	294	66
17	339	**336**	**336**	154
18	149	174	138	83
19	18	159	91	138
20	77	**100**	95	**301**
21	294	26	26	**199**
22	216	29	137	184
23	363	139	76	187
24	5	127	50	189
25	74	51	29	116

* Without mesoscale RUC or MM5

Tables 6a and 6b are colored coded as follows:

- blue bold font – Julian Day common to actual concentrations and to model-estimated concentrations using RUC and MM5
- black bold font – Julian Day common to model-estimated concentrations using RUC, MM5 and NWS-only meteorological data
- color shading – Julian Day common to actual concentrations and to all three types of meteorological data

Table 6b. Julian Days (2002) of 25 Highest Sulfur Dioxide Concentrations

TRNP-South Unit

Julian Day 1 = January 1, Julian Day 365 = December 31

Rank		Model-estimated using:		
1=highest, etc.	Actual	RUC	MM5	NWS Obs *
1	248	**78**	**337**	248
2	297	250	249	64
3	28	136	**126**	249
4	141	233	**128**	197
5	283	64	**73**	**73**
6	26	67	67	**78**
7	29	26	293	**39**
8	49	**73**	233	**337**
9	247	**126**	**82**	233
10	64	249	**78**	270
11	235	**173**	270	117
12	241	248	26	250
13	292	**337**	137	29
14	236	336	198	118
15	243	106	**301**	199
16	244	117	64	137
17	249	270	250	189
18	250	125	**173**	293
19	238	198	117	26
20	**173**	160	248	50
21	237	**82**	216	**126**
22	240	**39**	29	**82**
23	239	**128**	**39**	238
24	242	**301**	318	**301**
25	251	50	127	**128**

* Without mesoscale RUC or MM5

Seven of the 11 days from Julian Day 73 (March 14) through Julian Day 83 (March 24) are among the largest 25 actual sulfur dioxide 24-hour concentrations in the NU (see table 6a). Fifteen of the 17 days from Julian Day 235 (August 23) through Julian Day 251 (September 9) are among the largest 25 concentrations in the SU (see table 6b).

Pairing of model-estimated 24-hour sulfur dioxide concentrations day-to-day throughout all 365 days during a year with actual (monitored) 24-hour averaged concentrations improves insight into model accuracy. The matched daily pairs were rank sorted by actual concentrations and model-estimated concentrations, respectively. In this manner, one subset (containing the 25 highest actual concentrations with time-paired model-estimated concentrations) and another subset (containing the 25 highest model-estimated concentrations with time-paired actual concentrations) were extracted from the 365 days of paired concentrations. These subsets are not independent.

Temporal correlation is one metric of model performance. [118] Correlation within the two paired-in-time sets can be statistically expressed using Pearson correlation coefficients (PCC), as shown in table 7. No numeric standard of acceptable correlation is known; but any negative coefficients would be unacceptable.

Table 7. Correlation (PCC) Between SO2 Model-estimated * and Actual Concentrations

25 highest matched-in-time 24-hour pairs ranked by:

	Actual Concentrations			Model-estimated Concentrations		
	RUC	MM5	NWS-only	RUC	MM5	NWS-only
rural Hannover	**0.49**	0.18	0.17	**0.53**	0.11	0.06
TRNP--NU	0.66	0.69	0.08	0.74	0.49	-0.26
TRNP--SU	0.11	0.12	0.35	-0.23	-0.23	0.39

* SO2 background concentration for model-estimated SO2 concentrations is 0 ug/m3.

In sum, the temporal correlation in sulfur dioxide model estimated-concentrations matched with paired-in-time actual concentrations at the rural Hannover monitoring site is much better when using RUC data than when using MM5 or NWS data without RUC or MM5. Nevertheless, temporal correlation in matched pairs is poor. So model-estimated concentrations should not be used for tabulating estimates of changes in concentrations after PSD baseline using the paired-in-space-and-time method.

[118] See *Performance Measures and Standards for Air Quality Simulation Models*, EPA-450/4-79-032, October 1979, Office of Air Quality Planning and Standards, RTP, NC. Pages VI-10, VI-16 and VI-21.

5.8 Calmet–Calpuff estimated SO_2 concentrations include significant error and bias.

(See section 5.5.) Prior Calmet–Calpuff performance accuracy tests compare model-estimated sulfur dioxide concentrations paired in place but unpaired in time (e.g., day to day). [119] The NDDH has suggested better methods are needed for model results accuracy tests. [120]

(See section 5.7) Pairing of model-estimated 24-hour sulfur dioxide concentrations day-to-day throughout all 365 days during a year with actual 24-hour averaged concentrations was completed, and two subsets of matched paired-in-time concentrations were created.

Two descriptors of model accuracy are Mean Error (ME) and Mean Bias (MB). ME is the mean of absolute differences between matched daily pairs of model-estimated concentrations and actual concentrations. MB is the mean of those differences. Each difference is the model-estimated concentration deviation over (+) or under (−) the actual concentration occurring that same day.

$$ME = \overline{|(C_M - C_A)|} \qquad MB = \overline{(C_M - C_A)}$$

The "M" and "A" subscripts indicate values for modeled and actual concentrations, respectively. Mean bias cannot exceed mean error. In these formula, model-estimated concentrations are not background adjusted.

Scatter plots of the deviation of model-estimated sulfur dioxide 24-hour concentrations from actual concentrations (bias) are illustrated in figures 24, 25 and A7. Major sources contribute far more than oil and gas production flares and treaters to the model-estimated concentrations (see Attachment C).

The 25 pairs in each of the two subsets of matched paired-in-time concentrations were used to calculate ME and MB. The mean error and mean bias results, shown in table 8 and summarized below, assume that the background concentration for sulfur dioxide is 0 ug/m3.

When sorting matched daily pairs by actual concentration:

- ✗ ME analyses indicate models are, on average, overestimating or underestimating actual concentrations in TRNP in the range of 2.5 to 5.0 ug/m3.
- ✗ MB analyses indicate models are, on average, underestimating actual concentrations in TRNP in the range of 1.4 to 3.9 ug/m3.

[119] See, for example, Addendum C to Exhibit 158, sections 4, 5, and 6.

[120] Id., section 10. See also Addendum I to Exhibit 158, section 11.4.

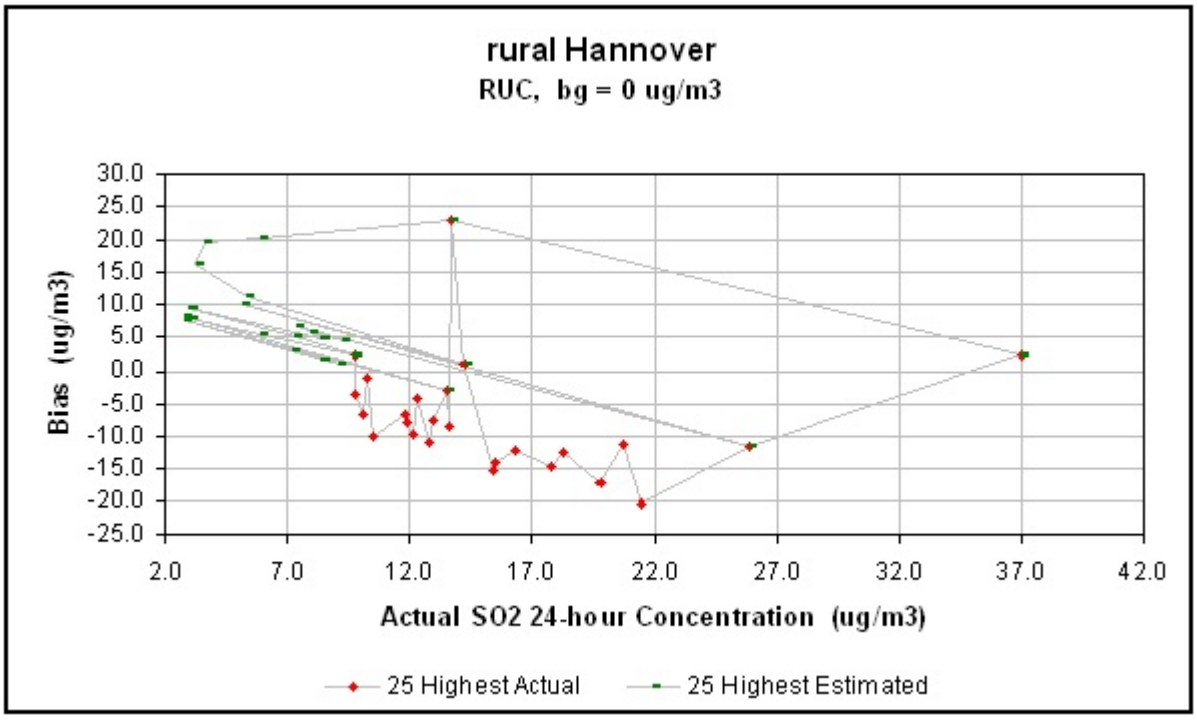


Figure 24. Bias in Model-estimated Sulfur Dioxide 24-hour Concentrations Time Paired with Actual Concentrations

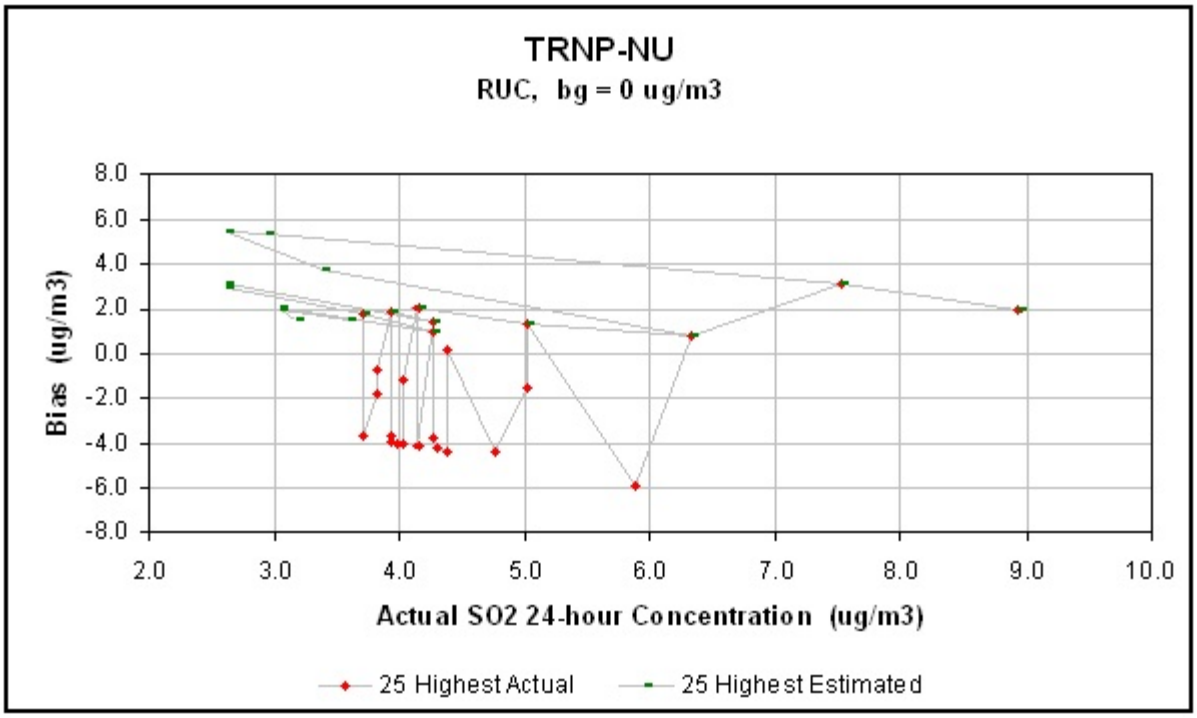


Figure 25. Bias in Model-estimated Sulfur Dioxide 24-hour Concentrations Time Paired with Actual Concentrations

When sorting matched daily pairs by model-estimated concentration:

✗ ME analyses indicates models are, on average, overestimating or underestimating actual sulfur dioxide concentrations in TRNP in the range of 1.7 to 3.2 ug/m3.

✗ MB analyses indicates models are, on average, overestimating actual sulfur dioxide concentrations in TRNP in the range of 0.5 to 2.7 ug/m3.

Table 8. Error and Bias in Model-estimated * Sulfur Dioxide 24-hour Concentrations Matched in Time with Actual Concentrations

	Largest 25 Actual Concentrations			Largest 25 Model-estimated Concentrations		
	RUC	MM5	NWS-only	RUC	MM5	NWS-only
rural Hannover						
ME (ug/m3)	9.6	10.3	9.8	**7.8**	11.3	13.0
MB (ug/m3)	-7.3	-6.6	-6.8	**6.6**	8.8	10.4
TRNP-NU						
ME (ug/m3)	2.7	2.5	3.3	2.1	1.7	2.4
MB (ug/m3)	-1.4	-2.0	-2.7	2.1	1.4	1.9
TRNP-SU						
ME (ug/m3)	5.0	4.7	4.5	3.2	2.1	2.6
MB (ug/m3)	-3.6	-4.6	-3.9	2.7	0.5	1.9

* SO2 background concentration for model-estimated SO2 concentrations is 0 ug/m3.

A background concentration other than 0 ug/m3, such as the 1.5 ug/m3 used in the State's MOU Protocol, can be added to all ME and MB data in table 8.

In sum, error data indicate that model-estimated sulfur dioxide 24-hour concentrations at sites of sulfur dioxide monitors in TRNP deviate from actual concentrations in the range from 1.7 to 5.0 ug/m3. Bias data in table 8 indicate the models are underestimating (left side) or overestimating (right side) actual concentrations in the range of -4.6 to 2.7 ug/m3. The PSD class I 24-hour increment is 5 ug/m3. The bias data also suggest that the Calmet–Calpuff models are not skilled in estimating (a) actual sulfur dioxide pathways between sources and monitors and (b) actual concentrations. [121]

[121] See also section 4.1 in Addendum C to Exhibit 158.

The 25 matched-in-time pairs of the two subsets (section 5.7) were merged into a third subset, and duplicate pairs were eliminated. The remaining pairs, N, are shown in table 9. Mean normalized error (MNE) and mean normalized bias (MNB) were calculated as follows.

$$MNE = \overline{|\,[(C_M + bg) - C_A]\,| / C_A}\,, \text{ where bg = background concentration}$$

$$MNB = \overline{[(C_M + bg) - C_A] / C_A} = \overline{R - 1.0}\,, \text{ where } R = (C_M + bg)/C_A$$

Results are also shown in table 9. Mean normalized bias cannot exceed mean normalized error. When MNB is larger than 1.0, the average of R is larger than 2.0.

Table 9. Correlation, Error and Bias in Model-estimated * 24-hour Sulfur Dioxide Concentrations Matched in Time with Actual Concentrations

	RUC		MM5		NWS-only	
	N		N		N	
rural Hannover						
PCC	45	**0.25**	43	-0.15	44	-0.24
MNE		**1.21**		1.55		1.61
MNB		**0.70**		1.02		1.13
TRNP-NU						
PCC	41	0.21	40	0.18	43	-0.38
MNE		0.84		0.73		0.88
MNB		0.50		0.37		0.45
TRNP-SU						
PCC	44	-0.23	43	-0.47	43	-0.24
MNE		1.01		0.74		0.91
MNB		0.45		0.17		0.36

* SO2 background concentration for model-estimated SO2 concentrations is 1.5 ug/m3.

Use of RUC data results in: better correlation, less error and less bias in model-estimated sulfur dioxide 24-hour concentrations at the rural Hannover monitoring site; and better consistency in correlation and bias among the three monitoring sites (e.g., 0.70, 0.50 and 0.45).

In sum, the error data indicate that model-estimated sulfur dioxide 24-hour concentrations are deviating, on average, from actual 24-hour concentrations in the range of 73 to 161%. The bias data indicate that the model-estimated concentrations, on average, are larger than actual sulfur dioxide concentrations by 17 to 113%, or that models and inputs are overestimating actual concentrations.

5.9 The models do not emulate real pollutant trajectories during periods of light wind speed and wind direction reversals.

Data plotted in figures 18 through 25 illustrate sporadic behavior of differences between highest model-estimated concentrations and rank-order matched highest actual concentrations. In 1998, EPA concluded that:

> "[The] complex interaction of transport, vertical mixing, and dispersion have an effect on concentrations with respect to downwind distances in CALPUFF. Occasionally, the accumulation of mass released over several hours will be transported in such a manner that the combined effect is to produce sharp localized maxima [[122]] in simulated concentration values. The occurrence of such events is not predictable. ... Calm winds play a part in these events. These maxima seem to occur at most locations in the receptor network, at all downwind distances. [[123]] When they occur, they seem to affect in particular the results of the shorter averaging periods." [124]

An example of the complex surface and 900 millibar (mb) wind regimes is illustrated with raw RUC2 data in figures 26 and 27. Map references for class I areas and major source locations are south-north Highways 85 and 83 and west-east Interstate I-94. Surface wind speeds are calm to light (about 3 m/s or less), and surface air flows east to west from Highway 83 to Highway 85 where surface air flows south to north. Surface air in the area of I-94 also flows south to north. At the same time, air at 900 mb generally flows south to north at 5 m/s (11.2 miles per hour) and higher.

In sum, two multi-day episodes of higher sulfur dioxide concentrations occurred during year 2002. (See tables 6a and 6b [125].) Significant horizontal and vertical wind speed and direction changes occur on occasion in paths of pollutant transport trajectories between sources of sulfur dioxide, monitors of actual concentrations and PSD class I areas. (Note the inconsistent results for the TRNP-SU and the TRNP-NU in tables 5, 7, 8 and 9.)

---

[122] Because each 24-hour averaging period (calendar day) consists of 24 1-hour modeled concentrations, larger 1-hour sulfur dioxide modeled concentrations are averaged with lower 24-hour concentrations. Nevertheless, some ratios of modeled to observed concentrations were large. (See figure 6, table 6 and Appendix E in Addendum C to Exhibit 158.)

[123] On the mesoscale, calm winds can occur at locations between sites of NWS surface weather stations and between sources of sulfur dioxide and PSD class I areas.

[124] See Exhibit 125 titled *A comparison of CALPUFF with ISC3*. Dated December 1998 by EPA, Office of Air Quality Planning and Standards, Research Triangle Park. Publication No. EPA-454/R-98-020. Page 20. "CALPUFF was run in a mode that enabled ISC3-type meteorological data as input, and therefore winds are horizontally homogenous for each hour." Id., page 3.

[125] See also WLI report in Addendum D to Exhibit 158 titled *Synoptic Analysis of Episodic Easterly Wind Events in Central-Western North Dakota for the Years 2000 – 2002.*

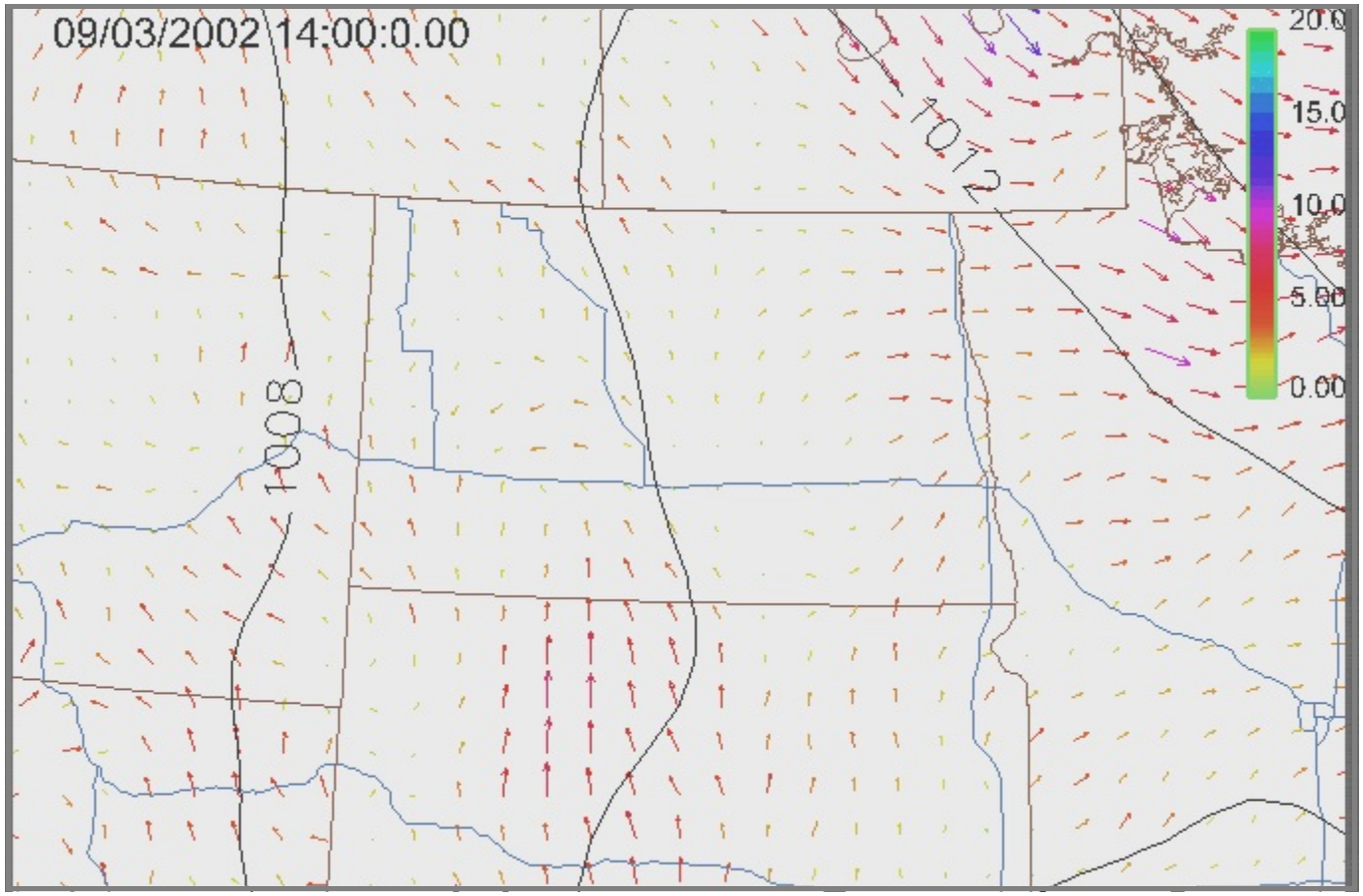


Figure 26. Surface pressure isobars (mb) and vector winds from raw RUC2 (40 km grid) data at 1400 UTC on September 3, 2002 (Julian Day 246). Color bar at upper right shows vector wind magnitude in m/s. The figure was taken from a surface vector wind animation prepared by WindLogics Inc (WLI).

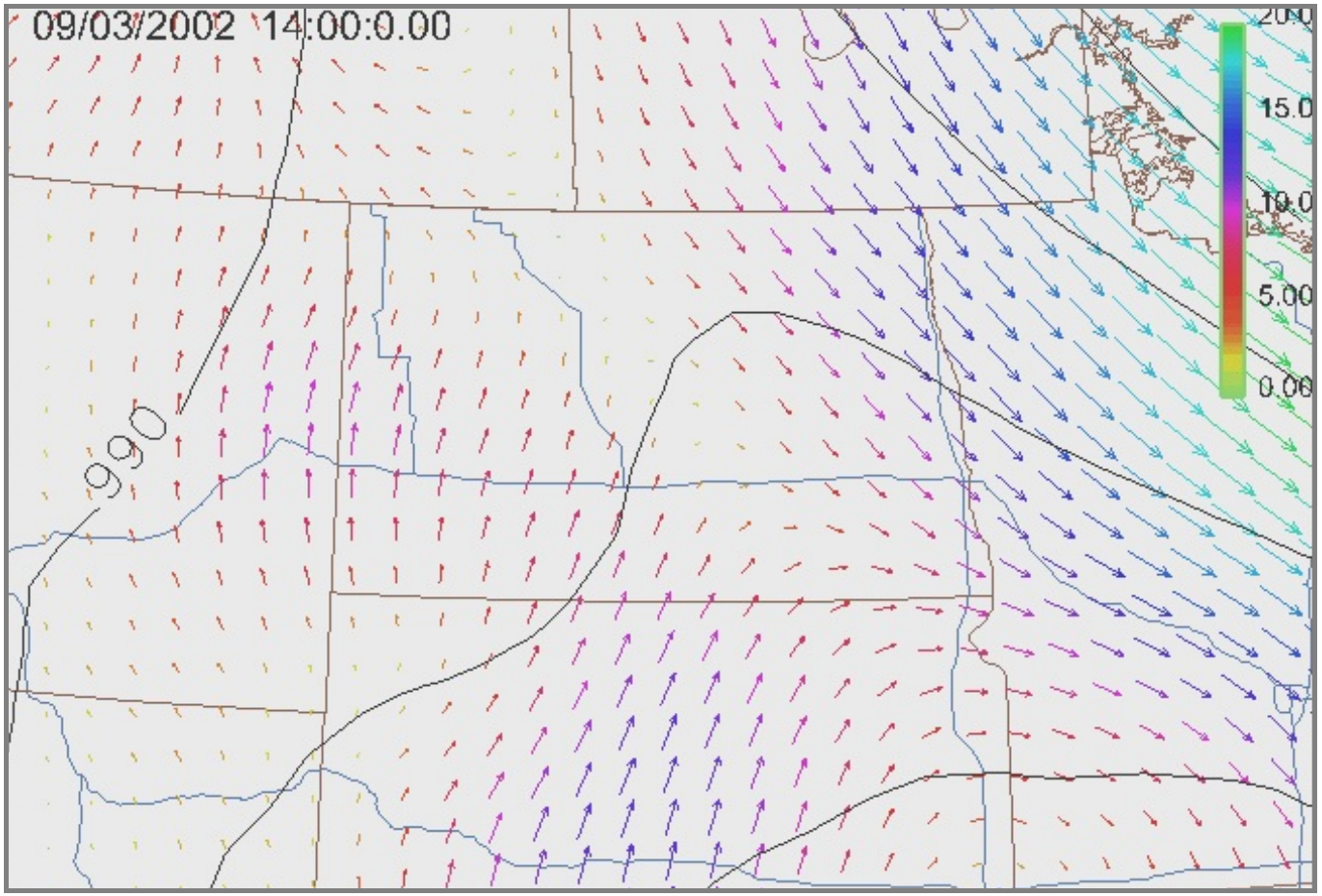


Figure 27. Geopotential height isopleths (m) and 900 mb vector winds from raw RUC2 (40 km grid) data at 1400 UTC on September 3, 2002 (Julian Day 246). Color bar at upper right shows vector wind magnitude in m/s. The figure was taken from a 900 mb vector wind animation prepared by WLI.

5.10 RUC mesoscale meteorological data are better than other meteorological data.

Accuracy analyses of RUC and MM5 data were completed. (1) RUC and MM5 data were compared to NWS observations and (2) Calmet output wind data, when using RUC or MM5 data as Calmet input, were compared to independent (not assimilated by RUC, MM5 or Calmet) wind-energy meteorological tower data. (See sections 5.2, 8.5 and 8.6)

✗ The RUC winds agree better with NWS surface observations of winds [126] and Calmet's output when using RUC as input agrees better with wind-energy meteorological tower observations of winds, [127] which provide implicit evidence that modeled pollutant transport trajectories using RUC data agree better with actual trajectories.

New Calpuff in-tandem-with-Calmet performance accuracy analyses were also completed. (See sections 5.3 through 5.8.)

✗ The largest model-estimated sulfur dioxide concentrations in TRNP when using RUC data are comparable to, or greater than, the largest estimated concentrations when substituting MM5 data for RUC data (figures 18, 20 and 21). The model-estimated concentrations when using RUC data are also greater than the largest actual sulfur dioxide concentrations. (See sections 5.3 through 5.6.)

✗ RUC data provide similar or better correlation between model-estimated sulfur dioxide concentrations and time-paired actual concentrations (table 7) when compared to correlation using MM5 data. And RUC data lead to lower normalized error and normalized bias at the rural Hannover monitoring site (tables 8 and 9) and more consistent normalized bias among the Hannover and TRNP sites (table 9). (See sections 5.7 and 5.8.)

Sources of sulfur dioxide are not co-located and are widely scattered in North Dakota. (See figure 1.) Six of the major sources were constructed after PSD baseline. (See table 2.) So meteorological data and Calmet and Calpuff modeling that best represents actual pollutant (model puff) transport and dispersion from each source to a model receptor is critical.

In sum, data in tables 5 through 9 and in figures 18 through 25 implicitly illustrate that meteorological data and the manner in which it was used with Calmet and with Calpuff contributed to differences between NDDH and EPA Region 8-estimated PSD sulfur dioxide short-term increment consumption.

---

[126] See Addendum D to Exhibit 158, which is titled *A Comparison of NOAA RUC Analysis Surface Winds and ADAS-Enhanced RUC Analysis Winds with Surface Observations*. See also Attachment 6 in Tab B of Volume 3, Exhibit 95 titled *A SSESCO [Windlogics] memorandum discussing an example of how interpolation between sounding data can underestimate wind speeds*.

[127] See Exhibit 95, Tab E, Volume 3, to Exhibit 95.

5.11 The NDDH's MOU PSD modeling and model performance accuracy tests used concurrent meteorology, source emissions and actual concentrations. (See discretionary options 5 and 6 per MOU.)

Results of air quality model performance accuracy analyses must be relevant to decisions of policy and decision makers who engage in management of air quality. Results also must be relevant to the disciplines of engineering and science.

> "In all applications of models an effort is encouraged to identify the reliability of the model estimates for that particular area and to determine the magnitude and sources of error associated with the use of the model. The analyst is responsible for recognizing and quantifying limitations in the accuracy, precision and sensitivity of the procedure. Information that might be useful to the decision maker in recognizing the seriousness of potential violations includes such model accuracy estimates as accuracy of peak predictions, bias, noise correlation, frequency distribution, spatial extent of high concentration, etc. Both space/time pairing of estimates and measurements and unpaired comparisons are recommended." [128]

For many years, perhaps since 1978 or 1980, air quality modeling advocates and practitioners have embraced the use of five years of meteorological data in air quality modeling. (Exhibit 132, FN 128, section 9.3.1.2.(a).)

> "The model user should acquire enough meteorological data to ensure that worst-case meteorological conditions are adequately represented in the model results." (Id., section 9.3.1.1(a).)

Conceptually, such worst-case concentrations would occur only once during the multi-year period of meteorological data. This approach is analogous, e.g., to a 5-year storm event used in flood control management. Worst-case concentrations generally relate to acute exposure of humans when conducting NSR NAAQS analyses rather than to chronic exposure of AQRVs in PSD class I areas. [129]

In its MOU Protocol, the State used three years of RUC mesoscale meteorological data.

> "Less than five, but at least three, years of meteorological data (need not be consecutive) may be used if mesoscale meteorological fields are available ..." (Id., section 9.3.1.2(d).)

---

[128] See Exhibit 132, section 10.1.3(b).

[129] FLMs have not identified or established acute exposure levels for AQRVs in North Dakota's PSD class I areas. The FLM initial certifications of no adverse impact occurred when actual sulfur dioxide concentrations in TRNP-NU were the highest ever (1982).

Often in modeling, the years of meteorological data did not coincide with the actual emissions used as input. For example, EPA Region 8's 2003 draft modeling used 1990 through 1994 meteorological data with estimates of increment-affecting sulfur dioxide 2000-01 emissions. [130] Prior to the EPA and State MOU, the NDDH had also used 1990 through 1994 meteorological data with current-period (2000–01) and PSD baseline emissions. [131]

Actual ambient concentrations represent known impact on air quality at the site of the monitor due to actual stack-top emissions. When inaccuracies in model-estimated concentrations are not illustrated with performance accuracy testing, policy and decision makers are inadequately informed and modeling practitioners are deviating from appropriate practice in engineering and science disciplines.

The NDDH, for example, demonstrated that sulfur dioxide emissions declined after 2000-01. [132] Matching 2002 actual emissions, instead of 2000-01 actual emissions per the MOU Protocol, with 2002 mesoscale meteorology reduced the inaccuracy (i.e, ratios) of model-estimated concentrations to actual concentrations about 20%. [133] So use of 2000-01 emissions instead of the lower emissions in 2002 distorts model-estimated sulfur dioxide deterioration from PSD baseline through year 2002 and PSD increment consumption. And model-estimated sulfur dioxide concentrations adjusted with a background concentration for 2002 continued to be larger than actual sulfur dioxide concentrations; that is, mean ratios were larger than 1.0, which is equivalent to mean normalized bias larger than 0.0.

In sum, the modeling completed by the NDDH pursuant to the State's MOU Protocol used concurrent meteorology and actual emissions (within the two years required by rule definition). And model performance accuracy tests and results described in the State's MOU Protocol Results Report [134] and in sections 5.2 through 5.8 of this paper used sulfur dioxide actual emission rates for the years 2002-2003, meteorology for the year 2002, and actual concentrations for the year 2002.

---

[130] See Exhibit 84, page 10.

[131] See Exhibit 81, section 2.2.

[132] Reasons for replacing 2000-01 sulfur dioxide emissions per the State's MOU Protocol with 2002-03 emissions are discussed on pages 23 through 26 in section 4.0 of Exhibit 158.

[133] See section 6.0 in Addendum C to Exhibit 158. Several ratios of the 25-highest, model-estimated 3-hour sulfur dioxide concentrations to 25-highest, 3-hour actual concentrations (when using 2002 meteorological data and 2000-01 averaged emission rates) were larger than 2. The ratios decreased to less than 2 when using 2002 meteorology and 2002 emission rates. (Id., Appendix E.) EPA has indicated that ratios of less than 2 (but larger than 1) are often quoted as acceptable. (See section 10.1.2 in Appendix W to 40 CFR Part 51.)

[134] See sections 5.0 and 6.0 in Addendum C to Exhibit 158.

5.12 In sum, model performance metrics for temporal correlation, error and bias indicate poor model performance.

When assessing the accuracy of model-estimated concentrations, "known" results of sulfur dioxide actually emitted, transported and dispersed by actually occurring meteorology are needed. The "known" results are typically concentrations actually measured and recorded.

When ambient sulfur dioxide concentrations are larger than the monitor's lower detection level, properly calibrated and operated monitors provide reliable actual sulfur dioxide hourly concentrations. The highest 25 24-hour average sulfur dioxide concentrations during the year 2002 at the rural Hannover, the NU of TRNP and the SU of TRNP monitoring sites are shown in table 10. When 24 consecutive hours of observed hourly concentrations are averaged, some hourly concentrations are less than the monitors' lowest reliable calibrated level (LRCL). When this occurred, one-half of the LRCL or 1 ppb (2.62 ug/m3) was used as the hourly concentration. This practice can cause inaccuracy in 24-hour sulfur dioxide actual concentrations, especially at the sites of the monitors in TRNP where sulfur dioxide actual concentrations are almost always low.

Analysis of Calpuff in-tandem-with-Calmet performance metrics in sections 5.3 through 5.8:

- ✗ Used only year-2002 meteorology. Pollutant trajectories for and dispersion of emitted sulfur dioxide vary from year to year, because meteorology varies from year to year. So statistical data for mean normalized bias (section 5.5), temporal correlation (section 5.7) and mean absolute error and mean bias (section 5.8) would likely vary from year to year.

- ✗ Used "actual emissions" as an average of total annual emissions during operating hours for years 2002-03. The performance metrics data in these sections quantify inaccuracy in model results in context for decision making in administration of the PSD program.

- ✗ Included the two sources operating under sulfur dioxide PSD class I increment variances. (See section 6.3.)

Sulfur dioxide monitors located in TRNP and at rural Dunn Center are located at greater distances from most larger sources of sulfur dioxide than the monitor at rural Hannover. The rural Hannover monitoring site is located about 20 to 40 km southeast, south, southwest and west of several power plants that emit sulfur dioxide. But the rural Dunn Center site is generally west of these sources, and the TRNP–NU and TRNP–SU sites are further west and more than 90 km downwind from most major sources.

So model performance metrics should indicate better agreement between sulfur dioxide model-estimated and actual concentrations at the site of the monitor at rural Hannover. As shown in table 7, the temporal correlation between model-estimated and actual concentrations is best (higher coefficient) when using the RUC mesoscale meteorological data. And as shown in

tables 5 and 8, error and bias in model-estimated concentrations are best (lower) when using the RUC mesoscale meteorological data.

✗ The implementing regulations' definition for "actual emissions" does not include an option to use sulfur dioxide hourly CEM emissions paired with hourly meteorology. However, pairing hourly CEM emissions with hourly RUC meteorology does not reduce error and bias in model-estimated concentrations (see Attachment B). Apparently, the atmosphere's stochastic nature influences pollutant transport and dispersion and, thus, performance metrics (compare figures 24, 25 and A3). Major sources contribute far more than oil and gas production flares and treaters to the model-estimated concentrations (see Attachment C).

At sites of sulfur dioxide monitors in TRNP:

✗ Error in modeled pollutant transport, dispersion and depletion is retained or increases as distances increase downwind of sources. The distances between most major sources and TRNP and the LNWA are greater than 90 km. So the stochastic nature of the atmosphere and imperfect model algorithms and model input data would not improve temporal correlation or reduce error and bias for model-estimated sulfur dioxide 24-hour concentrations as distances downwind of sources increase. (Compare performance metrics in tables 5, 7 and 9.)

✗ Model-estimated sulfur dioxide concentrations should not be used for tabulating estimates of changes in concentrations after PSD baseline using EPA's paired-in-time-and-space method, because:

- ✓ The temporal correlation in sulfur dioxide model-estimated 24-hour concentrations paired-in-time with actual concentrations is generally poor, since PCC are less than 0.25 (see table 9).
- ✓ Error in model-estimated sulfur dioxide 24-hour concentrations paired in time with actual concentrations ranges from 1.7 to 5.0 ug/m3 (see table 8), and bias ranges from -4.6 to 2.7 ug/m3; the PSD class I 24-hour increment is 5 ug/m3.
- ✓ There is no compelling reason to conclude that model performance metric data when using PSD baseline emissions and baseline meteorology would be better (i.e., larger temporal correlation coefficient, lower error and lower bias) than metric data when using current (2002-03) emissions and current meteorology.
- ✓ Meteorology, such as wind velocity and mixing height, is four dimensional – x, y, z, and t, and source locations changed after PSD baseline – some were shut down and others started up. So model errors in baseline sources' pollutant plume/puff transport and dispersion preceding and during a specific day (using current-period meteorology) likely do not match and cancel errors in current sources' pollutant plume/puff transport and dispersion.

Table 10. Number of Valid 1-hour SO2 Concentrations Less Than 2 ppb (5.24 ug/m3).

rural Hannover				TRNP--NU				TRNP--SU			
Julian Day	Actual	Lhrs	Vhrs	Julian Day	Actual	Lhrs	Vhrs	Julian Day	Actual	Lhrs	Vhrs
205	37.02	11	24	73	8.95	12	24	248	12.44	0	24
211	25.89	13	24	66	7.53	12	24	297	8.30	15	24
70	21.46	8	21	83	6.33	8	24	28	7.97	12	24
51	20.75	13	24	80	5.89	1	24	141	7.31	11	24
81	19.75	2	24	39	5.02	18	24	283	6.99	16	24
49	18.34	16	24	95	5.02	16	24	26	6.88	17	24
230	17.79	6	24	78	4.76	14	22	29	6.44	5	24
22	16.27	13	24	29	4.37	11	24	49	6.22	9	24
82	15.51	11	24	92	4.37	11	24	247	6.22	0	24
160	15.38	14	23	302	4.29	14	22	64	5.89	12	24
72	14.20	7	24	69	4.26	17	24	235	5.79	0	24
136	13.76	14	24	293	4.26	16	24	241	5.68	0	24
39	13.65	17	24	337	4.26	13	24	292	5.68	11	24
137	13.55	15	24	79	4.15	14	24	236	5.46	0	24
95	13.00	9	24	116	4.15	17	24	243	5.46	0	24
23	12.79	9	24	153	4.04	17	24	244	5.46	0	24
199	12.34	13	24	339	4.04	20	24	249	5.46	1	24
158	12.13	7	24	149	3.99	20	23	250	5.46	0	24
105	11.89	15	24	18	3.93	17	24	238	5.35	0	24
204	11.79	11	22	77	3.93	19	24	173	5.24	12	24
358	10.48	17	24	294	3.93	18	24	237	5.24	0	24
312	10.27	11	24	216	3.82	14	24	240	5.24	0	22
52	10.03	15	24	363	3.82	18	24	239	5.13	1	24
79	9.72	13	24	5	3.71	17	24	242	5.13	1	24
97	9.72	18	24	74	3.71	19	24	251	5.13	1	24

"Actual" is the average of all valid 1-hour SO2 concentrations in ug/m3.
Vhrs = # of hours of valid 1-hour SO2 concentrations Lhrs = # of Vhrs < LRCL LRCL = 2 ppb

Part 6

Calculating Increment Consumption

Using Modeled and Monitored Concentrations

6.1 The PSD "baseline concentration" is a single value for the PSD baseline period.

The federal CAA's NAAQS and PSD provisions for air quality deterioration, such as increment consumption and "baseline concentration", are based, respectively, upon the highest concentration or highest of second-highest (HSH) concentrations at air quality monitors or at model receptors.

The CAA at §§ 163(a) and (b)(1) states:

> (a) ... In the case of any maximum allowable increase (except an allowable increase specified under section 165(d)(2)(C)(iv)) for a pollutant based on concentrations permitted under national ambient air quality standards for any period other than an annual period, such regulations shall permit such maximum allowable increase to be exceeded during one such period per year.
> (b)(1) For any class I area, the maximum allowable increase in concentrations of sulfur dioxide and particulate matter ***over the baseline concentration*** of such pollutants shall not exceed the following amounts:
>
> | | Maximum allowable increase [Micrograms per cubic meter] |
> |---|---|
> | ... | |
> | Sulfur dioxide: | |
> | Annual arithmetic mean | 2 |
> | Twenty-four-hour maximum | 5 |
> | Three-hour maximum | 25 (emphasis added) |

The PSD increments do not stand alone; "Total ambient concentrations [or maximum allowable increases] ... consist of two components, baseline concentration and increment concentration." [135]

"'Baseline concentration' means that ***ambient concentration*** level which exists ***in the baseline area*** at the time of the applicable minor source baseline date. ..." [136] (emphasis and italics added) The baseline concentration for PSD is due to emissions during the PSD baseline period.

When using monitoring data (actual concentrations) to establish PSD baseline, EPA stated:

> "At a minimum, the data should be presented in a summary format showing the highest and highest, second highest concentrations for pollutants with short-term standards and the appropriate long-term average associated with each standard.

---

[135] See EPA's *Prevention of Significant Deterioration Workshop Manual* dated October 1980 at page I-C-3. See also Addendum H to Exhibit 158, page 3, and Addendum A to Exhibit 158.

[136] See Exhibit 57, page 20. See also Addendum G to Exhibit 158, pages 12 through 15.

> These concentrations effectively describe the existing ambient concentrations within the impact area attributable to actual emissions from existing sources." [137]

The highest and the highest, second highest monitored concentrations are single values from the baseline monitoring data collected with a network of two or more monitors. Twenty-four hour increment added to (over) the 24-hour baseline concentration establishes the maximum allowable ambient concentration, not to be exceeded by current 24-hour concentrations (at any location) within the PSD class I area more than once per year. [138] (See also section 6.3.)

Recently, EPA has stated:

> "EPA's Guideline on Air Quality Models requires that for PSD modeling, 'sequential modeling must demonstrate that allowable increments are not exceeded *temporally* and spatially, i.e., for all receptors for each time period throughout the year(s).' This means that to determine compliance with the PSD increment, one should determine whether the net change in increment consuming [sic] emissions since the baseline date has resulted [in changes] in pollutant concentrations exceeding the PSD increment at any specific time (temporal) and location (spatial) in the ***current*** year." [139] (emphasis and italics added to "current")

Neither the paragraph quoted above nor EPA's Guideline on Air Quality Models, which is Appendix W to 40 CFR Part 51, [140] include the term "baseline concentration".

EPA Region 8 did not model sulfur dioxide PSD baseline-period emissions; so it could not determine sulfur dioxide 24-hour and 3-hour baseline concentrations for model receptors or for each class I area. Its modeling of sulfur dioxide increment-affecting emissions reduces to 365 24-hour and 2,920 3-hour model-estimated changes in concentrations at each model receptor in a PSD class I area. Each model-estimated change implicitly includes a pseudo baseline concentration of magnitude that lies between the background concentration (without impact by baseline emissions) and larger values (due to impact by baseline emissions).

In sum, EPA's paired-in-space-and-time method of modeling for increment consumption (i.e., modeling increment-affecting emissions) do not match the provisions of the CAA's §§ 163 (a) and (b)(1); but instead its method exceeds those provisions. (See section 6.2.)

---

[137] See EPA's *Prevention of Significant Deterioration Workshop Manual* dated October 1980 at page I-C-23. See also CAA § 169(4).

[138] Id., pages I-C-33 and I-C-34.

[139] See Exhibit 136, which is a memorandum from Warren D. Peters, EPA's OAQPS, to Kevin Golden, EPA Region 8, dated October 16, 2002.

[140] See Exhibit 132.

6.2 The State's MOU Protocol included time-matching model estimates of concentrations using PSD baseline and current-period emissions. (See discretionary option 5 per MOU and section 8.8.)

In North Dakota, major sources of sulfur dioxide are widely separated and located at wide-ranging distances from the state's PSD class I areas. Some major sources were shut down and other sources started up after PSD baseline. There are 365 sequential 24-hour periods (days) and 2,920 sequential 3-hour periods during a non-leap year. So some 24-hour and 3-hour concentrations of sulfur dioxide at sites of monitoring instruments after the PSD baseline date could have decreased and other 24-hour and 3-hour concentrations increased.

The State's MOU Protocol included two methods to estimate short-term sulfur dioxide deterioration (increasing ambient concentrations) or increment consumption after PSD baseline. [141] The paired-in-space-only method incorporates a baseline concentration, while the paired-in-space-and-time method does not. The NDDH calculated deterioration using both methods and provided results in its MOU Protocol Results Report. Both methods provide temporal (all short-term periods throughout the current year so as to include changes in meteorology) and spatial (all model receptors [142]) increment exceedance analysis. (Exhibit 136, FN 139.)

Under the paired-in-space-only method, a short-term (24-hour and 3-hour) baseline concentration can be determined for ***each model receptor*** in a PSD class I area as the SH of sequential short-term estimated concentrations during the year by modeling PSD baseline sulfur dioxide emissions. [143] Sequential short-term estimated concentrations throughout the year are also determined at each model receptor by modeling current-period sulfur dioxide emissions.

- Deterioration at a receptor over that ***receptor's*** baseline concentration that exceeds a PSD class I increment is an exceedance. [144] In this approach, there possibly could be 365 24-hour or 2,920 3-hour exceedances at each receptor. However, plumes change direction; so, in general, exceedances are not meteorologically independent. [145]

---

[141] See section 6.2 in Addendum B to Exhibit 158. See also section 8.0 in Addendum C to Exhibit 158.

[142] See section 6.1 in Addendum B to Exhibit 158.

[143] Id.

[144] The CAA's § 163(a) and implementing regulations allow one exceedance of short-term increments per year. Section 9.3 in Addendum C to Exhibit 158 illustrates, for year 2002, the spatial variation of increases over the baseline concentration among receptors in the South and North Units of TRNP.

[145] See Addendum C to Exhibit 158, section 9.2 and figures 13 and 14. See also Exhibit 22 titled *Air Quality Effects Analysis of Basin Electric Power Cooperative Antelope Valley Station Unit 3*, section 3.4.3.1, and Exhibit 23 titled *Air Quality Effects Analysis of Warren Petroleum Natural Gas Processing Plant (Expansion)*, section 3.4.3.

Under a third method, the "baseline concentration" is the highest (worst-case) of SH estimated 24-hour or 3-hour concentrations among all receptors in the PSD class I area when modeling PSD baseline emissions. Under a fourth method, receptor concentrations for each 24-hour or 3-hour period throughout the year are averaged for the receptor network; the "baseline concentration" is the SH average, 24-hour or 3-hour, during the year. [146] These methods were not used. These methods conform to the CAA and PSD implementing regulations, where "baseline concentration" as defined is that ambient concentration level in the ***baseline area*** at the time of the minor source baseline date. (See section 6.1.) These methods also provide temporal (all short-term periods throughout the current year) and spatial (all model receptors) increment exceedance analysis.

- Deterioration at a receptor over a ***class I area's*** "baseline concentration" that exceeds a PSD class I increment is an exceedance. And in these approaches, there possibly could also be 365 24-hour or 2,920 3-hour exceedances at each receptor. Again, plumes change direction; so, in general, exceedances are not meteorologically independent.

In sum, several model receptors are used in each of the state's class I areas, except the TRNP Elkhorn Ranch. The paired-in-space-and-time method implicitly uses a pseudo baseline concentration for each short-term period throughout the year at each receptor. (See section 6.1.) The paired-in-space-only method explicitly uses the SH short-term concentration during the year at a receptor as that receptor's baseline concentration. So, the paired-in-space-and-time method is likely to demonstrate more exceedances than the paired-in-space-only method. Both methods are more stringent than application of "baseline concentration" as defined by the CAA and implementing regulations.

The CAA's §§ 163(a) and (b)(1) detract from the use of other methods for tabulating changes in air quality. [147] For example, the change could be the difference between averages of the 10 highest model-estimated concentrations at a receptor using both current and baseline emissions.

> "Sole reliance on a single value [i.e., SH or HSH] out of the multiplicity of available model outputs narrows both the definition of the problem and the perspective of the decision-maker." (FN 147, page 13.)

One or more additional methods of tabulating model-estimated concentrations and estimated changes in air quality would expand the perspective of modelers and decision makers. The NDDH has suggested some additional methods. [148]

---

[146] No reference to an air quality receptor is included in sections 163(a) and (b)(1) of the CAA.

[147] See *Workshop Summary Report: Role of Atmospheric Models in Regulatory Decision-Making*, EPA A-80-46, II-M-6. Pages 23 and 24.

[148] See Addendum I to Exhibit 158, section 10.

6.3 The State's MOU Protocol included, as well as excluded, sources granted variances to the SO_2 PSD 24-hour increment per FLM certifications of no-adverse impact.

The CAA at §§ 165(d)(2)(C)(iii) and (iv) states:

> "(iii) In any case where the owner or operator of such facility demonstrates to the satisfaction of the Federal Land Manager, and the Federal Land Manager so certifies, that the emissions from such facility will have no adverse impact on the air quality-related values of such lands (including visibility) notwithstanding the fact that the change in air quality resulting from emissions from such facility will cause or contribute to concentrations which exceed the maximum allowable increases for class I areas, the State may issue a permit."
> "(iv) In the case of a permit issued pursuant to clause (iii), such facility shall comply with such emission limitations under such permit as may be necessary to assure that emissions of sulfur oxides and particulates from such facility, will not cause or contribute to concentrations of such pollutant which exceed the following maximum allowable increases ***over the baseline concentration*** for such pollutants:
>
> Maximum allowable increase
> [Micrograms per cubic meter]
>
> ...
> Sulfur dioxide:
> Annual arithmetic mean 20
> Twenty-four-hour maximum 91
> Three-hour maximum . . . . . 325" (emphasis added)

The State has completed a detailed legal analysis of administrative implementation of PSD class I variances. [149]

> "North Dakota is the only state that has existing and active Federal Land Manager (FLM) alternate Class I variances issued under CAA § 165(d)(2)(C)(iii)." [150] "The two facilities currently operating under these FLM alternate variances are the Dakota Gasification Company (DGC) that operates near Beulah, North Dakota, in the central part of the state, and the Little Knife Gas (Little Knife) plant, that operates in the west-central part of the state." [151]

In sum, the State's MOU Protocol and Protocol Results Report included estimates for sulfur dioxide increment consumption with and without the sulfur dioxide emissions of the two sources granted PSD variances per FLM certifications of no-adverse impact.

---

[149] See Addendum H in Exhibit 158 titled *The PSD Variance Issue in North Dakota*. See also Exhibits 13, 127, and 155, paragraph 6.6.

[150] Id., page ii.

[151] Id., page ii.

6.4 Atmospheric boundary layer meteorology, influenced at the surface by terrain, affect ground-level concentrations.

The highest, SH (HSH) model-estimated sulfur dioxide 24-hour concentrations in the NU of TRNP generally occurred in the southwest panhandle of the area. [152] The HSH in the SU generally occurred in the southeast corner of this area. [153] (See figures 28 and 29.)

- The range (difference between highest and lowest) of SH model-estimated sulfur dioxide 24-hour concentrations for receptors in the SU of TRNP for 2002 is 4.9 ug/m3, and the range in the NU for 2002 was 4.7. [154] However, data in table 8 (page 66, RUC) indicate that the mean error in model-estimated sulfur dioxide 24-hour concentrations at the monitoring site in the SU of TRNP ranges between 3.2 and 5.0 ug/m3 and in the NU between 2.1 and 2.7 ug/m3.

Analyses of measured winds reveal terrain influences on wind direction and speed.

- The dominant wind directions at the TRNP-SU monitoring site are inconsistent with the dominant wind directions at the rural Dunn Center site. [155]

- Data in figure A3 (page 122) illustrates that the average wind speeds during larger actual concentrations at the TRNP-NU monitoring site are less than the average speeds at the rural Hanover and TRNP-SU monitoring sites.

Analyses of actual ambient sulfur dioxide concentrations reveal meteorological influences on concentrations.

- Data in table 10 (page 76) illustrate that pollutant plumes are transient at monitoring sites. The dwell time of plumes containing larger concentrations over a monitoring site is typically about one-half day. Plumes sweep left (counter clockwise) or right (clockwise) as wind directions change during the passage of weather systems.

- Data in figures 8 (page 10) and A2 (page 121) reveal a stochastic relationship between concentrations and surface wind speed at respective monitoring sites.

In sum, pollutant dilution is proportional to wind speed at stack top; but factors in addition to wind speed and terrain dominate subsequent pollutant transport and dispersion leading to on-the-ground actual concentrations.

---

[152] Addendum C to Exhibit 158, pages 67, 68 and 69.

[153] Id., pages 20, 73 and 74.

[154] Id., pages 69 and 74, respectively.

[155] See Appendix A, Addendum C to Exhibit 158.

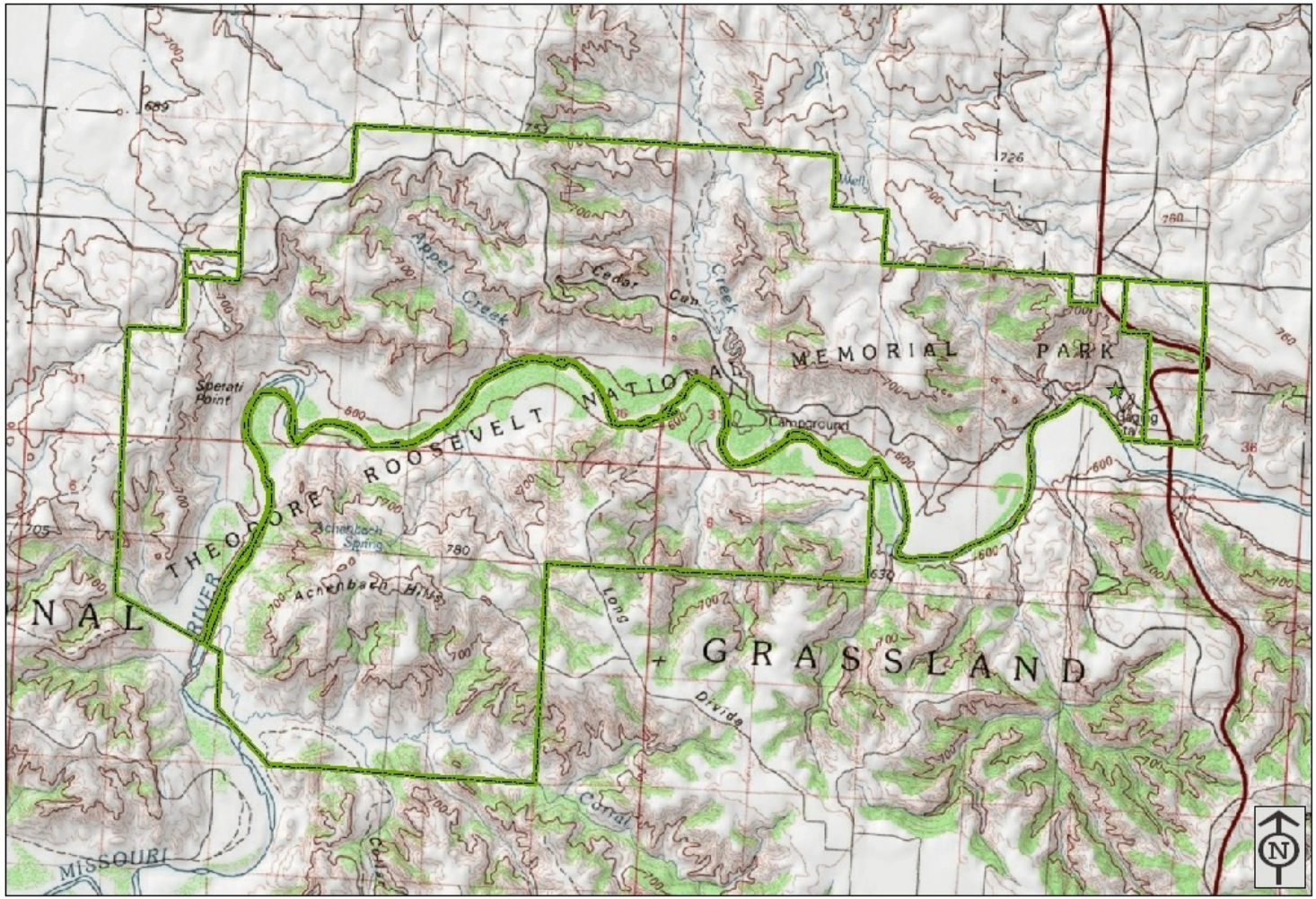


Figure 28. Topographic map of North Unit of TRNP

Figure 29. Topographic map of the South Unit of TRNP

6.5 The spatial densities of model receptors in the state's PSD class areas are adequate.

The Calmet and Calpuff models were used to estimate on-the-ground sulfur dioxide 1-hour concentrations for networks of model receptors in the state's PSD class I areas. Sequential blocks of 3 1-hour and 24 1-hour concentrations were averaged for estimating 3-hour and 24-hour (daily) concentrations, respectively. An example follows in figure 30.

A model receptor was also located at the sites of monitors in TRNP, and model-estimated concentrations at these receptors were used for model performance accuracy analyses.

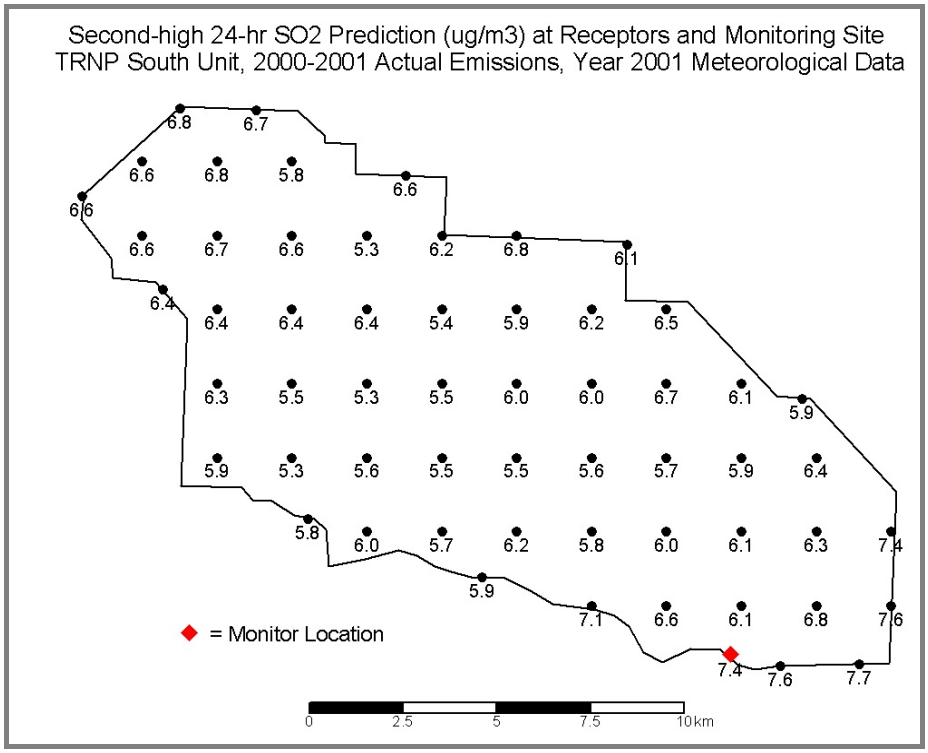


Figure 30.

- Distributions of estimated concentrations at neighboring model receptors are likely not statistically independent. When this is so, useful accuracy in modeled concentrations is not improved by reducing the spacing of receptors in receptor networks.

- The day/hour timing for an exceedance of a short-term increment at one receptor in a class I area can overlap the day/hour for an exceedance at another due to changing wind directions. (See sections 6.2 and 6.4.) So total exceedances among receptors do not express possible violations of the increments. However, total exceedances are dependent on the number of extreme weather events that impact the area. [156]

[156] See Addendum C to Exhibit 158, section 9.2.

6.6 Subsection 165(e) of the CAA requires air quality monitoring.

The CAA at §§165(e)(1) and (2) states:

> "(1) The review provided for in subsection (a) of this section [i.e., NAAQS and PSD NSR review] ***shall be preceded by an analysis*** in accordance with regulations of the Administrator, promulgated under this subsection, which may be conducted by the State (or any general purpose unit of local government) or by the major emitting facility applying for such permit, ***of the ambient air quality at the proposed site and in areas which may be affected by emissions from such facility*** for each pollutant subject to regulation under this chapter which will be emitted from such facility." (emphasis and italics added)
>
> "(2) Effective one year after August 7, 1977, the analysis required by this subsection ***shall include continuous air quality monitoring data*** **gathered for purposes of determining whether emissions from such facility will exceed the maximum allowable increases or the maximum allowable concentration permitted under this part**. ..." (emphasis and italics added)

Actual sulfur dioxide concentrations during PSD baseline are needed to establish the "baseline concentration" so that changes in actual concentrations in subsequent years "over the baseline concentration" can be calculated and compared to the maximum allowable increases (the increments). Prior to 1980, no valid or reliable actual concentrations for sulfur dioxide are available for locations in TRNP-NU and TRNP-SU. The year 1980 was the first complete year of actual concentrations obtained in each unit of the park. And the monitoring site in the LNWA was installed during 2003.

Beginning in 1975, the NDDH has continuously used air quality models to assess new source PSD increment consumption. [157] Since 1980, there has been no useful analysis of existing post-1979 actual sulfur dioxide concentrations for estimating sulfur dioxide short-term baseline concentrations during PSD baseline years. [158] EPA Region 8, for example, discussed an implicit relationship between local oil and gas emissions of sulfur dioxide and actual concentrations in TRNP during the 1980s and a probable trend for actual concentrations preceding 1980 back-in-time to PSD baseline years. [159] (Compare figures 4 and A1 to figures 6 and 7.) It did not, however, use annual oil production data from the 1970s and 1980s to estimate sulfur dioxide short-term baseline concentrations.

---

[157] See Exhibit 133, table 7.

[158] Testimony during the NDDH's May 2002 hearing included graphs of SH concentrations extrapolated back into PSD baseline years, but the extrapolated concentrations for those years exceeded SH concentrations during the early and mid 1980s in contradiction to historical in-state coal and oil production. See also letter under Tab D in Vol. 6 to Exhibit 95, pages 1 and 2.

[159] See Exhibit 57, pages 9 and 10.

6.7 EPA considers actual concentrations reliable for tracking NAAQS compliance and for model accuracy tests. (For tracking NAAQS compliance, see Addendum F to Exhibit 158; and for model accuracy tests, see sections 3.6, 3.8 and 5.1.)

Regarding use of actual concentrations (monitoring data) for tracking amounts of consumed PSD increment or available PSD increment, EPA Region 8 states:

> "EPA generally considers monitoring data unreliable for determining how much of the increment has been used up. Several factors are worthy of note here. ... Third, it is not practical to have monitors in all locations where elevated concentrations of pollutants may threaten PSD increment. Fourth, models have the advantage of being able to predict pollutant and PSD increment concentrations at locations where siting of monitors may not be possible. Fifth, due to the lack of an adequate number of monitors in the early years of the PSD program (during the time period the baseline was established), [sic] ***if the program were to rely on monitoring it would make calculating baseline (and other aspects of the PSD program) virtually unworkable***. Finally, monitoring data collected at a single location is not representative of concentrations that may occur at other nearby Class I receptors because SO2 concentrations can vary greatly over small distances. For these reasons, EPA believes that the assessment of available increment will normally be accomplished through an accounting procedure whereby modeling results will be used to keep track of the available increment." [160] (references omitted and emphasis and italics added)

No single factor given by Region 8 negates using available actual sulfur dioxide concentrations obtained with the one monitor in the South Unit of TRNP and the one monitor in the North Unit for tracking trends in sulfur dioxide concentrations, as shown in figures 6 and 7, at the sites of these monitors. And Region 8's stated fifth factor does not negate back-in-time extrapolations of actual concentrations at sites of monitors in both units of TRNP to estimate baseline concentrations at those sites. [161]

---

[160] See Exhibit 57, pages 8 and 9. Counterpoints to each of the factors are omitted from the discussion details of this paper. (See, for example, Exhibit 95 titled *Post-Hearing Comments to the North Dakota Department of Health*, and *Comments on EPA Region 8's May 2003 Modeling of Class I SO2 Increment Consumption*, pages 13–17.) And in an apparent 1980 contradiction, EPA described the role of monitoring data to establish existing ambient concentrations attributable to actual emissions from existing sources. See *Prevention of Significant Deterioration – Workshop Manual*, OAQPS, RTP, October 1980, pages I-C-22 and 23.

[161] "However, the monitored data show a large decrease in SO2 concentrations at Theodore Roosevelt National Park-North Unit in the two years preceding the peak concentrations measured in 1982. If that trend had continued back to the 1977 time period, coincident with the reduced oil production, concentrations in the 1976 to 1977 baseline period would have been lower than those monitored in 1980, or even in current years." Id., page 9.

6.8 Actual concentrations can be used to track trends in annual second-highest 24-hour concentrations.

Prior to 1980, no valid or reliable actual concentrations for sulfur dioxide are available for locations in TRNP-NU and TRNP-SU. The year 1980 was the first complete year of actual concentrations obtained with monitors in each unit of the park. Figure 6 shows the annual SH of 24-hour averaged actual sulfur dioxide concentrations from data obtained with the monitors.

Baseline for PSD is a two-year period. The "baseline concentration" represents an adjusted time period that considers dates for commencement of construction and operation of major stationary sources. [162]

- One coal-fired utility began construction after the major source baseline date but was not in operation on the minor source baseline date. This source, Coal Creek Station's Unit 1, first began operating during May 1979. [163]

- The PSD minor source baseline date for western North Dakota, which is December 19, 1977, was triggered by a completed application for the Little Knife Gas Plant. This plant began operating during July 1978. [164]

- Unit 2 of the Leland Olds plant had start-up problems that extended into year 1976.

- The NDDH determined that the two years representative of normal operations for some sources were 1977-78 or 1978-79. [165] (See section 4.4.)

In sum, the best-fit years for PSD baseline are 1977-78 when tracking changes in actual SH sulfur dioxide concentrations.

Actual SH sulfur dioxide concentrations for PSD baseline years could be estimated by examining year-to-year trends in: (1) on-line power production capacity of electric utility boilers, as shown in figure 31, and (2) oil production, as shown in figures 2, 4, A1 and 11. However, the coal-fired utility boilers were not continuously operated at rated maximum heat-input capacity during PSD baseline and subsequent years. [166]

---

[162] See Addendum H to Exhibit 158, page 3. See also: section 2.3 in Addendum G to Exhibit 158, 45 FR 52713–52715, and EPA's *Prevention of Significant Deterioration – Workshop Manual*, OAQPS, RTP, October 1980, page I-C-3.

[163] See Addendum B, Appendix C, to Exhibit 158.

[164] Id., See also Exhibit 24.

[165] See Addendum B, Appendix D, to Exhibit 158.

[166] See Exhibit 83, pages 37 through 89. See also Exhibit 82, pages 51 through 53.

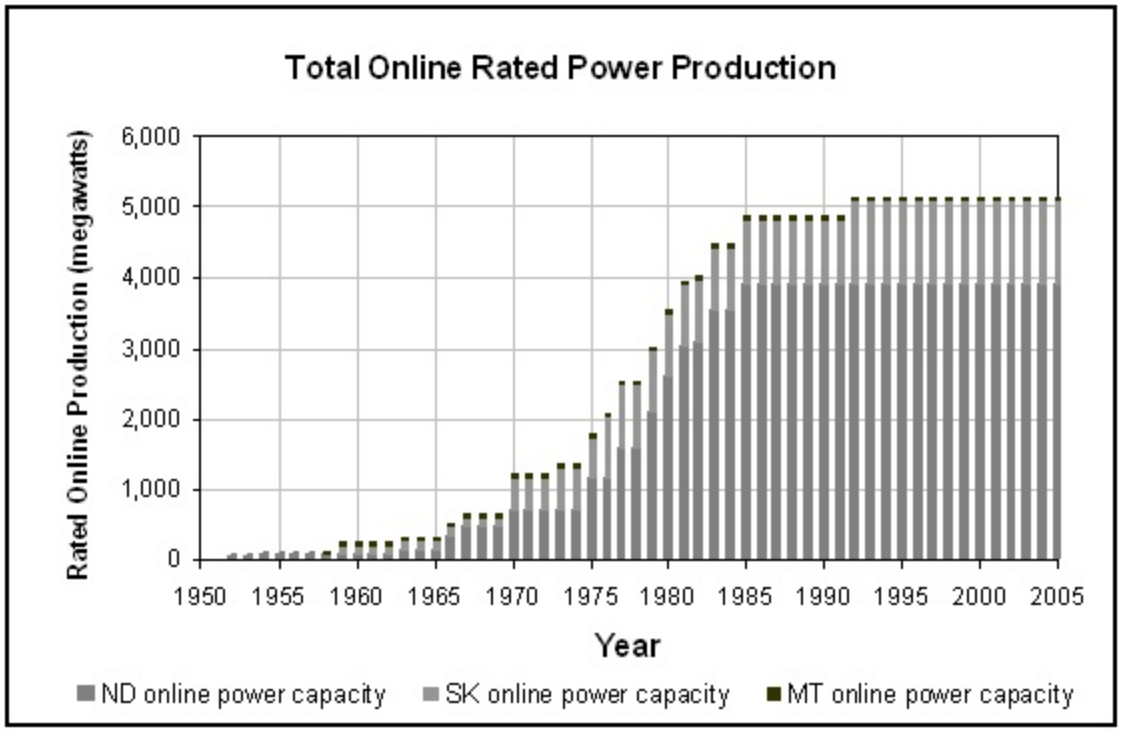


Figure 31. (Plants near Sidney MT and near Regina SK are included.)

The NDDH does not have data representing annual consumption of coal by the state's electric utilities prior to 1981. Historical coal production data for North Dakota are available from the U.S. Bureau of Mines,[167] as shown in figure 32. The amount of coal mined in North Dakota during 1977-78 was ~70% of the amount mined during 1980.

The annual oil production in McKenzie County during 1978 was ~74% of the amount produced during 1980, in Billings County ~21% and statewide ~61%, as shown in figures A1, 4 and 11. Use of coal and oil production data to extrapolate a 1980 SH actual sulfur dioxide concentration back to a 1977-78 baseline requires some assumptions: (1) all mined coal was fired in utility boilers, and annual emissions of boilers were equally proportional to amounts of coal mined, (2) annual emissions of oil-field flares and treaters were also equally proportional to oil production, which is conservative since sour natural gas in oil and gas flaring declined as oil production matured (see sections 4.8 and 4.9), and (3) worst-case pollutant transport and dispersion meteorology during 1977-78 was similar to worst-case meteorology during 1980.

So the SH actual sulfur dioxide concentration during 1977-78 might have been ~70% of the SH concentration during 1980. The SH sulfur dioxide concentrations during 1980 for the TRNP-NU and the TRNP-SU were 21.0 ug/m3 and 7.9 ug/m3, respectively. And estimates of the

[167] Source: http://www.eia.doe.gov/cneaf/coal/statepro/imagemap/nd.ht. 1890-1978 – U.S. Bureau of Mines file data, including State-level statistics not published in Bureau of Mines Mineral Industry Yearbooks. 1979-present – Energy Information Administration, Annual Coal Report (2001) and precedent reports (Coal Industry Annual, 1993-2000; and Coal Production, 1979-1992).

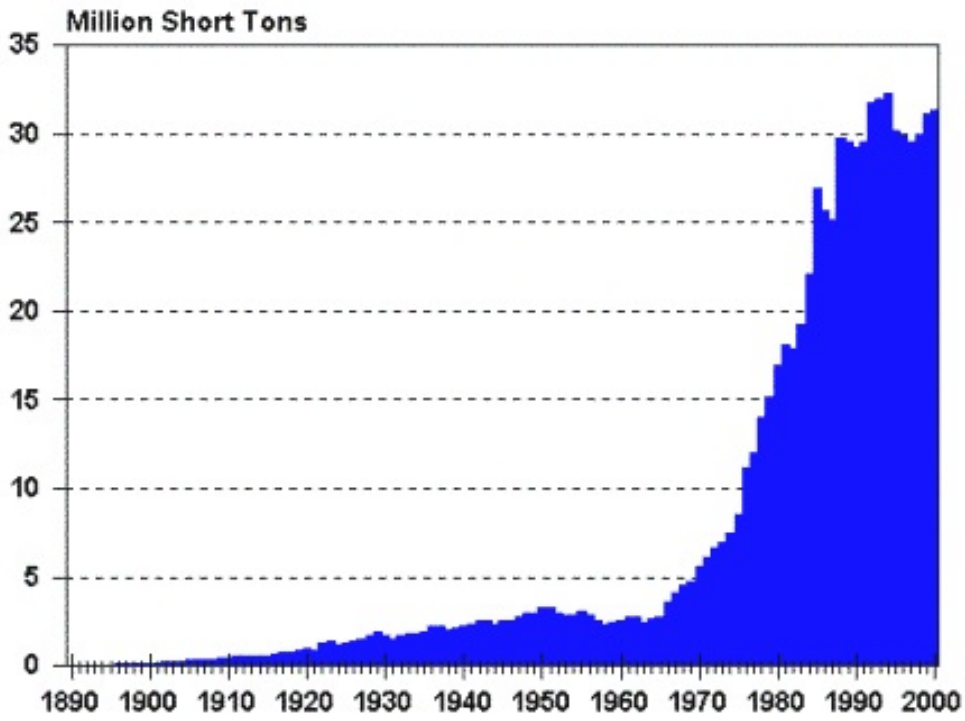


Figure 32. Historical Coal Production for North Dakota

sulfur dioxide "baseline concentration"for the TRNP-NU and the TRNP-SU are 14.7 ug/m3 and 5.5 ug/m3, respectively.

The estimated baseline concentrations were subtracted from annual SH [168] actual sulfur dioxide concentrations (see figure 6) for years 1980 through 2004 to illustrate changes in annual SH concentrations "over [or under] the baseline concentration," as shown in figure 33.

The dashed line at 5 ug/m3 represents the PSD class I sulfur dioxide 24-hour increment over the baseline concentration. The solid line at 91 ug/m3 represents the PSD class I sulfur dioxide alternate 24-hour increment over the baseline concentration. [169]

In sum, estimated changes "over [or under] the baseline concentration" in annual SH 24-hour actual concentrations since 1980:

➙ Were sometimes larger than the sulfur dioxide PSD class I 24-hour increment of 5 ug/m3 in both units of TRNP, but since 1998 have been less than 5 ug/m3.

[168] The CAA provides that the SO2 PSD class I short-term (24-hour and 3-hour) increments for sulfur dioxide can be exceeded once per year.

[169] See Addendum H to Exhibit 158, 11 through 17.

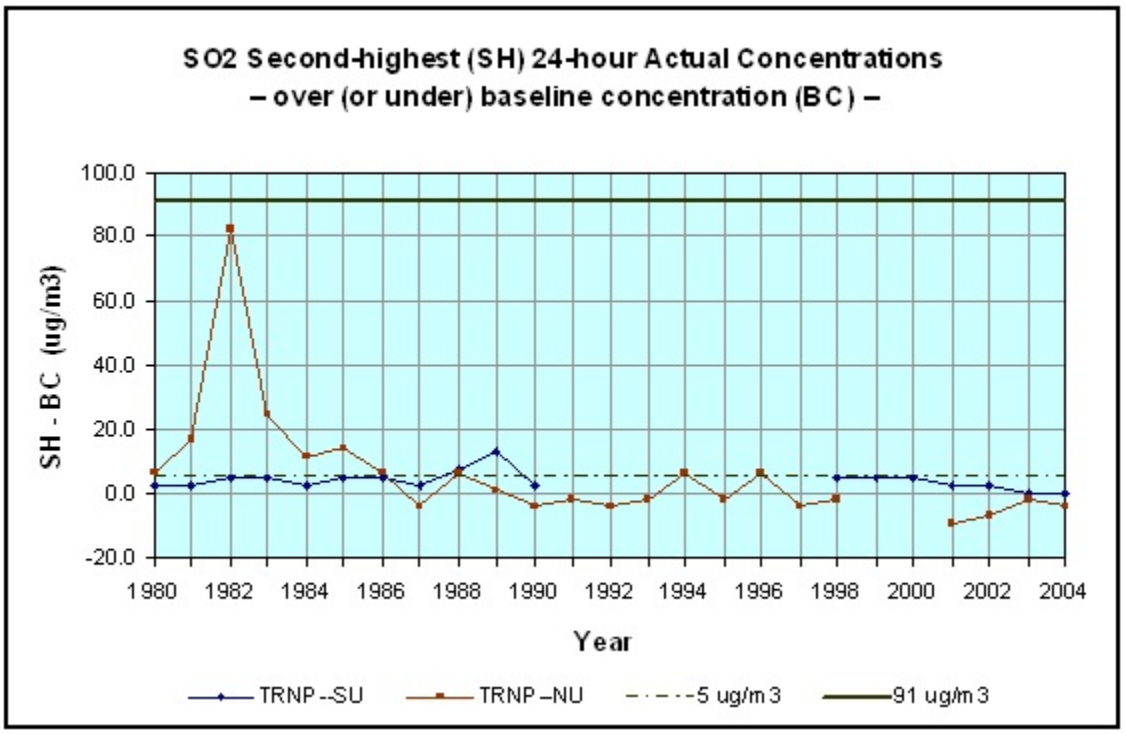


Figure 33.

- Have never exceeded the sulfur dioxide PSD class I alternate 24-hour increment of 91 ug/m3.

- Illustrate implementation of a paired-in-space (place)-only analysis of actual concentrations.

The U.S. Department of Interior published final certifications of no adverse impact on AQRVs in the TRNP and the LNWA as follows: [170]

- 47 FR 41480-01 dated September 20, 1982
- 49 FR 38197-02 dated September 27, 1984
- 50 FR 7658-04 dated February 25, 1985
- 58 FR 13639-01 dated March 12, 1993

"[T]he [1982] FLM certification of no adverse impact occurred at the time that the highest monitored concentrations of SO2 ever recorded in TRNP were occurring. Since then, the air quality has improved significantly from the 1982 levels." [171]

[170] See Exhibit 13, which contains copies of USDOI federal register notices. See also Addendum H to Exhibit 158, pages 9 through 19.

[171] See Addendum H to Exhibit 158, page 14.

Part 7

Final – *North Dakota's*

SO2 PSD Air Quality Modeling Report

7.1 The State's MOU Protocol contains several adjustments to the NDDH's draft 2003 protocol.

Following the State's periodic review hearing process in 2002 and 2003, the State Health Officer issued Findings and Conclusions dated September 8, 2003. [172] Section 10.0 of the Findings and Conclusions states:

> "The evidence shows that the Department [the NDDH] has engaged EPA since the last hearing [2002] to resolve outstanding issues through a written agreement [i.e., Exhibit 88] and through continued negotiations. It should continue to do so subsequent to this order."

Consequently, a new MOU (Addendum A to Exhibit 158) and a new modeling protocol (Addendum B to Exhibit 158) were negotiated by the State and EPA during the winter of 2003-04 and the spring of 2004. (See section P.1.) During oral negotiations, EPA recommended several adjustments to the NDDH's draft 2003 modeling protocol. (See section 3.10.) **The NDDH adopted all EPA (OAQPS and Region 8) comments on drafts of its MOU Protocol.** [173] A summary of these adjustments follows.

A. During negotiation of the MOU, the NDDH agreed to:

1. Adjust the sulfur content of feed coal for utility boilers to the two years of normal source operations. (See: State Health Officer's September-8-2003 Findings and Conclusions, section 10.1; MOU, part 1.3; and Addendum I, section 6.1.A.) This change, in combination with adjustments to sulfur dioxide emission factors for power plants, increased PSD increment-affecting sulfur dioxide emissions about 11,080 tons per year.

2. Assign the same two-year period for normal operations to all emitting units at a power plant. (See section 6.1.C in Addendum I and Appendix D to Addendum B.)

3. Provide an additional written technical description of the ADAS-enhanced RUC data and a statistical comparison of this data with NWS surface wind data. Two reports were completed for the NDDH via contract with WindLogics, Inc., St. Paul, MN. (See Addendum D.)

   ✓ *RUC Analysis-based CALMET Meteorological Data for the State of North Dakota* dated August 24, 2004.

   ✓ *A Comparison of NOAA RUC Analysis Surface Winds and ADAS-Enhanced RUC Analysis Winds with Surface Observations* dated August 27, 2004.

---

[172] See Addendum E to Exhibit 158.

[173] See Addendum I to Exhibit 158, section 2.

Subsequently, the NDDH arranged a workshop, hosted by WindLogics, Inc., in St. Paul during July 2004, where the ADAS-enhanced RUC data and the comparison analyses were orally presented to representatives of the NDDH and EPA.

4. Use sulfur dioxide emission factors for lignite-fired power plants that are consistent with other data sources. Consequently, some sulfur dioxide emission factors were adjusted. (See part 1.3 in Addendum A, section 6.1.B(1 & 2) in Addendum I and footnote 12 on page 46 and table on page 47 (footnote "***" for Minnkota Unit 1) in Appendix E in Addendum B.)

B. During EPA review and comment on drafts of the MOU Protocol, the NDDH:

1. Clarified the role of model results sensitivity tests, which compare modeled outcome against modeled outcome when changing model inputs or model algorithms. (See section 1.3 (3rd paragraph) in Addendum B and section 2.2 in Addendum C.)

2. Completed an analysis of actual ambient sulfur dioxide 1-hour concentrations to determine an estimate of a background sulfur dioxide concentration for use in model results accuracy analyses. (See section 9 (3rd paragraph) in Addendum I and section 5.2 in Addendum B.)

3. Placed several model receptors at the perimeter of the NU and the SU of TRNP. (See section 5.2 in Addendum I and section 6.1 in Addendum B.)

4. Completed an additional, thorough review and analysis for source-specific sulfur dioxide emission factors for lignite fired power-plant boilers, since factors for both units of the Leland Olds Station were larger the EPA's AP-42 average factor of 32.3S. At 40S, all sulfur in combusted coal is emitted from power-plant stacks. (See section 6.1.B(3) in Addendum I and Appendix E to Addendum B.)

5. Evaluated concentrations of alkaline constituents in the coal and coal ash of one power plant, since these constituents, especially sodium oxide, are natural sulfur scrubbing agents that reduce amounts of emitted sulfur dioxide. (See section 6.2 in Addendum I and Appendix E in Addendum B.)

6. Edited and clarified discussion for (1) tabulating model-estimated sulfur dioxide concentrations to determine PSD class I sulfur dioxide increment consumption (see section 6.2 in Addendum B) and (2) normal source operations (see Appendix D to Addendum B).

In addition, the NDDH completed a thorough review of Calmet and Calpuff user input data for model control variables and adjusted several data settings or values. (See section 2.D through H and K, section 4.1 and section 5.1 in Addendum I.) EPA's OAQPS or Region 8 did not inquire or comment on these adjustments.

7.2 The State's MOU Protocol can be updated "after considering performance through model accuracy performance testing."

When the NDDH announced its April 2005 hearing, [174] it included public notice of a general policy pertaining to models and model inputs.

> "The Department is also aware of recent EPA-approved refinements to CALMET and CALPUFF, and will continue to consider these and other EPA-approved air quality models as they become available or are submitted to the Department in future PSD or New Source Review proceedings (within the constraints of its budget and resources). In addition, the Department desires to use more refined meteorological data, monitoring data, and emissions data, as those data become available, after considering performance through model accuracy testing. This is consistent with the State's general policy to use air quality monitoring to provide an empirical basis by which modeling techniques are verified and 'held to earth' by a continuing iterative process of confirmation and reassessment as better techniques, technically upgraded models, and more meteorological data, monitoring data, and emissions data become available for use in air quality assessments." (FN 174.)

This policy was confirmed in the State Health Officer's Findings and Conclusions following the second of two periodic review hearings regarding the status of SO2 PSD class I increment consumption. [175]

> "The policy is grounded in the testimony, evidence, and legal authority presented in the record which supports considering relevant air quality monitoring to provide an empirical basis by which the modeling techniques are verified and 'held to earth' ..." (Exhibit 155, FN 175, paragraph 6.5.)

> "The final modeling protocol (Addendum B) is adopted as a guideline pursuant to N.D.C.C. § 28-32-01(11)(k) for use in North Dakota's PSD/NSR program. Its use as a guideline for future PSD/NSR modeling is subject to the Department's general policy of using more refined meteorological data, monitoring data, and emissions data after considering performance through model accuracy testing." (Id.)

[174] See http://ndhealth.gov/AQ/Dockets/PSD/Notice%20of%20Hearing%20April%2019,%202005.pdf

[175] See Exhibit 155.

7.3 FINAL – *North Dakota's SO2 PSD Air Quality Modeling Report*

In August 2005, the State completed a CAA PSD periodic review of the status of attainment of PSD sulfur dioxide increments. On September 9, 2005, the NDDH forwarded to EPA: (1) a copy of State Health Officer's Determination signed September 7, 2005 (Exhibit 155), (2) the NDDH's Response to Comments pertaining to a hearing held April 19, 2005 (Exhibit 156) and (3) the State's SO2 PSD Air Quality Modeling Report (Exhibit 158). [176] The following information was provided with the report.

Glossary of acronyms and abbreviations.

Report addenda tab index.

A. North Dakota and U.S. EPA Memorandum of Understanding dated and signed February 24, 2004.

B. FINAL – *A proposed alternative air quality modeling protocol to examine the status of attainment of PSD Class I increments* dated August 18, 2005.

C. FINAL – *Results of air quality modeling to examine the status of attainment of PSD Class I sulfur dioxide increments* dated August 18, 2005.

D. *RUC Analysis-based CALMET Meteorological Data for the State of North Dakota* dated August 24, 2004. Prepared by WindLogics, Inc., St. Paul, MN.

A Comparison of NOAA RUC Analysis Surface Winds and ADAS-Enhanced RUC Analysis Winds with Surface Observations dated August 27, 2004. Prepared by WindLogics, Inc., St, Paul, MN.

Synoptic Analysis of Episodic Easterly Wind Events in Central-Western North Dakota for the Years 2000 – 2002 dated December 16, 2004. Prepared by WindLogics, Inc., St, Paul, MN.

E. August 8, 2002, and September 8, 2003, Orders of North Dakota State Health Officer Terry L. Dwelle.

F. FINAL – *Monitored Air Quality in North Dakota: A Summary* dated August 18, 2005.

G. FINAL – *A Summary of the Development of the Clean Air Act And Its Prevention of Significant Deterioration Provisions* dated August 18, 2005.

H. FINAL – *PSD Variance Issue in North Dakota* dated August 18, 2005.

I. FINAL – *Background Discussion of Model Input Data and Potential Refinements* dated August 18, 2005.

[176] See http://ndhealth.gov/AQ/Notices.htm

Table 11. Lists of modeling issues that occurred during negotiation of the MOU and the State's MOU Protocol
(The regional modelers issues appeared one week prior to EPA's verbal approval of the protocol.)

Regional Modelers' Issues	MOU	Locations in ND's Final SO2 PSD Report
I. monitoring & accuracy testing	III.3	overview: NDDH's report at section 3.4.3 methods & data: B 5., Appen. B, G & H // I 8, 9. results: C 1.2, 4., 5., 10.1, 10.2, 10.3, Appen. B, C & E
II. variances	II.1	overview: H // B 3.7 methods for inclusion & exclusion: B 1.4, 6.2 results: C 1.3, 1.4, 8.1
III. annual average rates (a.k.a. actual emissions)	I .2 & I.4	overview: NDDH's report at section 3.4, 3.5 methods & data: B 4., Appen. C, D, E & F // C 2.1, 6. // I 6.2, 11.2
IV. pairing in time...& space	I.5 & II.2	overview: NDDH's report at section 3.8, 3.5 methods: B 6.2 // I 8, 10. 11.5 results: C 1.2, 7., 8.1, Appen D, F, G & H
V. network averaging	I.5	method: I 8.
VI. baseline normal operations	I.3 & I.5	overview: NDDH's report at section 3.2 methods & data: B 4., Appen. D // C 2.1

Table 11 cont. Lists of modeling issues that occurred during negotiation of the MOU and the State's MOU Protocol

Additional EPA and State Issues ----------		
use of Calmet & Calpuff	I.1	overview: NDDH's report at section 3.1 implementation, data & rational: B 2.3, 3. // I 4.1, 5.1,11.1
meteorological data	I.6	overview: NDDH's report at section 3.6 description: B 2.1, Attach. 1 & 2 // D implementation: C 3. // I 3., 4.2, 11.3
model sensitivity	III.3	description: B 1.3 implementation: C 2.2 // I 11.6
background concentration	---	description: B 5.1, 5.2 implementation: C e.g., 5.2, 6.1, 10.3, Appen. E
model receptors	---	description: B 6.1 // I 5.2 implementation: C e.g., 5.1 & Appen. D
coal sulfur content	I.3	implementation: B Appen. D // I 6.1.A
minor source emissions	---	implementation: B 4.0, Appen. F
baseline concentration	I.5	overview: NDDH's report at section 3.8 implementation: B 6.2 // I 11.5 results: C 1.2, 8.1
emission factors	I.3	overview: NDDH's report at section 3.3 implementation: B 4.0, Appen. E // I 6.1.B, 6.2

Table 11 cont. Lists of modeling issues that occurred during negotiation of the MOU and the State's MOU Protocol

Other Information	
overview of the MOU Protocol:	B Exec. Sum. // C 1.1
brief history of modeling:	B 1.
changes to prior NDDH protocols & development of the MOU Protocol:	I 1., 2., 3., 4., 5., 6., 11.6
status of existing air quality	F

Part 8

Additional Details

8.1 Emphasis on national consistency in modeling

The EPA, the NPS and other FLMs use monitoring data to assess current conditions and trends in sulfur dioxide, sulfates, fine particulate, ozone and other pollutants. Their modelers, however, seem to prefer practices that give deference to modeling over monitoring and to absurd conservative choices in air quality modeling protocols. [177] This deference has occurred with no regard for inaccuracy in model results – in spite of emphasis on performance accuracy analyses since the early 1980s and a court decision, which emphasized that modeling techniques and results are to be verified through confirmation and reassessment.

EPA's continuing unduly conservative approach in using air quality models to assess new source emissions compliance with NAAQS and PSD increments, as exemplified by Region 8's draft 2002 and 2003 reports, is reflected in 1979 statements:

> "The NAAQS are written in quite specific terms and must ultimately be complied with. An argument can be made that to 'guarantee' such compliance, uncertainty in model predictions must be on the 'conservative' side. That is, the probability must be acceptably small that a control strategy designed [sic] based on model predictions will not actually achieve compliance. ... While the rationale has some potential usefulness, it implies the introduction of systematic bias into modeling results, something we would hope to avoid in a final choice of a [model] performance standard." [178]

> "By insisting that the model 'overpredict' peak concentrations, almost certainly we will select abatement strategies requiring more control than needed. [179]

Historically, federal agency air-quality modelers have often advocated national consistency in use of models and model input data. A primary reason for the early emphasis on consistency was to reduce the burden of oversight of modeling methods and interpretation of modeling results. In May 2005, Warren Peters, Director of OAQPS's Model Clearinghouse, writes in regard to future PSD/NSR modeling: "[T]here are new demands on old dispersion modeling

---

[177] For example, "we have no objection to your use of CEM data to determine a single emissions value that represents actual emissions patterns for each source, but we believe that you should use two consecutive years of CEM data to determine the maximum, or near maximum, emission rate just as you would if you were using permitted potential emissions. That single emissions value for each source would then be modeled over 5 years of meteorological data to identify expected increment violations under realistic conditions." (See Letter by John S. Seitz, OAQPS, dated December 12, 2001, to Francis J. Schwindt, NDDH.) See also "the Gulf Coast problem", which was described by EPA in 1980 at 45 FR 52681 and 52718 and is summarized in Exhibit 158 at section 3.4.2.

[178] See *Performance Measures and Standards for Air Quality Simulation Models*, EPA-450/4-79-032, October 1979, Office of Air Quality Planning and Standards, Research Triangle Park, NC. Page VI-25.

[179] Id., page D-37.

policies and practices – [we] must be confident of appropriate model applications and ensure national consistency and equity." [180]

EPA regional modelers released an internet email memorandum dated April 21, 2004 [181] that contains comments on an initial draft of the State's MOU Protocol dated March 9, 2004. This memorandum reflects deference for national consistency; it was dated one week prior to EPA's verbal approval of the protocol. Each of the email listed issues are addressed in the MOU or in other documents. (See sections 3.9 and 7.1.) In addition, there were other model protocol issues which arose during negotiation of the MOU and among comments on drafts of the State's MOU Protocol. (See table 11.)

8.2 Summary of EPA Region 8's flawed modeling

EPA Region 8's 2003 modeling protocol [182] deviated from the CAA, implementing regulations and court decisions as follows.

✗ (See section 4.6.) Region 8 used 90th percentiles of 24-hour averages of 2000-2001 sulfur dioxide hourly CEMS data as surrogate peak emission rates for several utility boilers. [183] Maximum, or peak, and ninetieth percentile rates are not among options in the regulatory definition for "actual emissions". [184] Actual emissions, PSD increment consumption and PSD baseline concentration are interrelated (see section 6.1). In a letter subsequent to the NDDH SIP periodic review and all modeling protocols described in this paper, EPA Region 8 states: "***The increment consumption analysis is based on 'actual emissions' (40 C.F.R. 52.166(b)(13))*** ..." [185] (emphasis and italics added) The correct citations for "baseline concentration" and "actual emissions" are 40 CFR § 51.166(b)(13) and 40 CFR §§ 51.166(b)(21)(i)&(ii), respectively.

✗ (See section 4.6.) For several sources, including some utility boilers, Region 8 used average annual sulfur dioxide emission rates computed as total annual emissions divided by 365 days (8,760 hours) rather than by operating hours, which results in lower emission rates since sources typically are not operated all hours during a year. Rates as averages during all 365 days are not among options in the regulatory definition for "actual emissions". (Exhibit 158, FN 184.)

---

[180] See http://cleanairinfo.com/modelingworkshop/presentations/Future_Modeling_Peters.pdf, page 13.

[181] See http://www.epa.gov/Region8/foia/ndair/ndmp.html This document is also Exhibit 140.

[182] See Exhibit 84.

[183] Id., pages 19, 20, 24, 27 and 28.

[184] See Exhibit 158, section 3.4.

[185] See letter from Richard R. Long, Director Air and Radiation Program, EPA Region 8, to Terry L. O'Clair, Director Division of Air Quality, NDDH, dated June 27, 2006.

✗ (See sections 6.1 and 8.7.) It modeled increment-consuming or expanding sulfur dioxide emissions of major sources as changes in emissions rates from PSD baseline to current baseline rather than modeling both PSD baseline and current sulfur dioxide emissions. [186] So it could not determine a sulfur dioxide 3-hour and 24-hour "baseline concentration" for each class I area, and it could not determine deterioration "over the baseline concentration."

✗ (See sections 6.1 and 6.2.) Region 8's modeling of increment-consuming or expanding sulfur dioxide emissions implicitly applies a baseline concentration for each of 365 sequential 24-hour periods and 2,920 sequential 3-hour periods throughout the year. This practice is more stringent than required by the definition for "baseline concentration" in the CAA and EPA's implementing regulations.

✗ (See sections 3.6 – 3.8 and 4.6.) It did not complete model performance accuracy tests using its chosen Calmet and Calpuff input data, including larger current-period (2000-01) sulfur dioxide emission rates. [187] It could not, since it modeled (a) increment-consuming or expanding emissions [188] and (b) 1990 through 1994 meteorology. Actual sulfur dioxide concentrations are due to sulfur dioxide actually emitted and dispersed by concurrent meteorology.

✗ (See section 6.3.) It included sources granted certifications of no adverse impact by FLMs when assessing the status of attainment of the CAA § 163 SO2 PSD class I increments. [189] However, these sources contribute to consumption of the CAA § 165 class I alternate increments.

EPA Region 8's 2003 modeling protocol deviated from its agency's guidance and its own precepts as follows.

✗ (See sections 5.6.) Region 8 indicated that "MM5 as currently configured tends to overestimate wind speeds, particularly in the lower levels of the atmosphere, unless restrained by actual observations" and that small values for Calmet variables R1 and R2 "replace all of the actual measured surface and upper air weather observations" with mesoscale meteorology (RUC or MM5). [190] As shown in section 5.2, Calmet output wind fields – when using NWS wind observations or MM5 wind data as Calmet input – underestimate, rather than overestimate, meteorological-tower wind speeds inside the NDDH modeling domain.

---

[186] See Exhibit 84, page 19 and pages 36, 37 and 38.

[187] Id., pages 14 and 15, figure 2-3.

[188] Id., page 19.

[189] Id., page 24.

[190] Id., page 12 and page 9, respectively.

✗ (See section 4.5.) It used EPA's estimated average sulfur dioxide emission factor for lignite-fired boilers of 30S rather than using 32.3S reflecting sodium oxide concentrations in North Dakota lignite or developing source-specific sulfur dioxide emission factors for those state utility boilers without emission control systems. [191] This oversight did not provide "apples-to-apples" comparable emission rates for both current years and PSD baseline years for these boilers. (This practice distorts PSD increment-affecting emissions and exaggerates increment consumption.)

✗ (See sections 4.2 and 4.6.) It used a statistical mix of 90th percentile and average daily sulfur dioxide emission rates. [192] This oversight did not provide "apples-to-apples" comparable sulfur dioxide emission rates for both current years and PSD baseline years among all sources; i.e., sources retired after PSD baseline, sources placed into operation after PSD baseline as well as sources operating during both time lines. [193] (This practice also distorts PSD increment-affecting emissions and exaggerates increment consumption.)

✗ (See section 4.6.) It included a sulfur dioxide emission rate for both the main stack and for the bypass stack at the Great Plains Synfuels Plant, which is a double counting and increased increment-affecting emissions by 855 lb/hr. (This practice also exaggerated PSD increment consumption.)

8.3 Calmet and Calpuff computer codes

The NDDH used older versions of Calmet and Calpuff per EPA oral approval on April 28, 2004; specifically, Calmet version 5.2 and Calpuff version 5.4 were used. [194] In its final report dated August 19, 2005, the State indicated that newer versions of Calmet and Calpuff would subsequently be used. [195] Newer versions of both models included error fixes, as well as upgrades in extrapolation of weather observations taken at NWS stations to the models' computational grid. [196]

---

[191] See Exhibit 84, pages 31, 32 and 33.

[192] Id., pages 19 through 34.

[193] See also Exhibit 17, letter by Richard R. Long, EPA Region 8, to Francis J. Schwindt, NDDH, dated June 25, 2001, paragraph 2(2).

[194] See Tab B to Exhibit 115, pages 7 and 13. See also paragraph I.1 in the MOU, which is Addendum A to Exhibit 158. Calmet version 5.5 and Calpuff version 5.7, which were approved by EPA effective April 15, 2003, were not used by the NDDH as indicated in sections 2.2 and 3.0 of Addendum B to Exhibit 158.

[195] See pages 27 and 28 in section 4.0 of Exhibit 158.

[196] See Exhibit 153, page 3-1.

Tests of the newer versions of models' effect on modeling results were completed. Test results indicate decreases or increases in model-estimated sulfur dioxide PSD class I 24-hour increment consumption and no exceedances of the increment. [197] The tests used sulfur dioxide emission rates during 2002-03 as in the State's MOU Protocol Results Report with year 2002 meteorology and rates during 2000-01 with 2000 and 2001 meteorology as in the State's MOU Protocol. [198]

During the winter months of 2004-05, the NDDH completed a limited test of the effects of using the April 2003 versions of Calmet and Calpuff on sulfur dioxide PSD class I 24-hour increment consumption. All inputs for the models followed the MOU Protocol except sulfur dioxide emissions. The results are shown in table 12. [199]

Since April 2003, additional Calmet and Calpuff Fortran-code error fixes and enhancements have been identified and applied. [200] Calmet version 5.53a and Calpuff version 5.711a were approved by EPA on June 15, 2006.

In sum, the newer April-2003 versions of the models caused change in model-estimated increment consumption in both units of TRNP. For example, the paired-in-space-only result (4.0 ug/m3) for the SU of TRNP is no longer larger than the paired-in-space-and time result (4.5 ug/m3), which is now, but was not previously, consistent with precepts for these methods.

8.4 New source reviews

PSD NSRs are bound to follow the State's MOU Protocol pursuant to the State Health Officer's Findings and Conclusions (Findings) from public hearings governed by North Dakota's Administrative Procedures Act. [201] The Findings also include a path-forward process for upgrading the State's protocol "after considering model performance through model accuracy testing." [202]

---

[197] Id., tables 3-1 through 3-4.

[198] Reasons for replacing 2000-01 sulfur dioxide emissions with 2002-03 emissions are discussed on pages 23 through 26 in section 4.0 of Exhibit 158.

[199] Data in table 12 for model versions 5.2/5.4 were taken from results presented in the tables in section 8 of Addendum C to Exhibit 158.

[200] See Model Change Bulletin A(040716) at http://src.com/calpuff/040716_MCB_a.txt

[201] See Findings at http://ndhealth.gov/AQ/Dockets/ModelingProtocolfinal/Findings%20and%20Determination%20of%20September%207,%202005.pdf This document is also Exhibit 155.

[202] Id., paragraph 6.5. See also paragraphs 1.a, 1.b and 1.d of Appendix W attached to 40 CFR Part 51. Appendix W is Exhibit 132.

Table 12. Model Influences on Model-estimated SO2 PSD class I 24-hour Increment Consumption (ug/m3)		
Results include sources having FLM CONAI		
Meteorology ?	2002	
SO2 Emissions ?	2002-03	
Calmet/Calpuff Versions ?	5.2 / 5.4	5.5 / 5.7
TRNP South Unit		
Paired in space & time	4.7	4.5
Paired in space only	4.9	4.0
TRNP Elkhorn Ranch		
Paired in space & time	0.7	0.7
Paired in space only	-8.0	-8.0
TRNP North Unit		
Paired in space & time	4.5	4.4
Paired in space only	-4.3	-4.2
Lostwood NWA		
Paired in space & time	2.5	2.5
Paired in space only	2.5	2.5

8.5 Background concentration for SO_2

Those sources located within the modeling domain, but not modeled, contribute to the background sulfur dioxide concentration. [203] These sources are mobile sources, all oil and gas production sources having sulfur dioxide emission rates less than 0.001 grams per second, and oil and gas production sources located beyond 50 kilometers from class I areas.

Other sources that are outside the domain also contribute to the background sulfur dioxide concentration. These sources are: mobile sources, heating plants and commodity processing plants in eastern North Dakota, sources in central and western Montana, and sources in Minnesota, South Dakota and southern Saskatchewan.

In its draft May 2003 modeling report, EPA Region 8 states:

> "The residence time of SO2 in the atmosphere is limited because SO2 is oxidized to particulate sulfate and is removed through deposition and other mechanisms.

[203] See section 5.0 to Addendum B to Exhibit 158.

> For this reason monitored concentrations from very distant sources of SO2 are likely negligible. ... EPA believes that a background value of 0.2 ppb is appropriate for the TRNP–South Unit." [204]

"The background concentration represents that sulfur dioxide in the air at the time when, and the places where, those sources in the emissions inventory emit sulfur dioxide." [205] Times of transport of sulfur dioxide from distant sources in the modeling domain to the state's PSD class I areas are likely less than 30 hours (150 km ÷ 5 km/hr). The residence time of SO2 in the cooler and arid climate of the plains of central North America is likely in the range of 50 hours, and perhaps longer.

Following EPA's advice when it provided comments on a draft of the State's MOU Protocol, the NDDH estimated the background concentrations by extrapolating hourly actual concentrations taken from the TRNP–SU, TRNP–NU and rural Hannover monitoring sites. [206] The result was an estimated background concentration for western North Dakota at 1.5 ug/m3 or 0.57 ppb. [207]

When tabulating changes in model-estimated concentrations from PSD baseline to current time using respective sulfur dioxide emissions inventories, the background concentration drops out and does not impact the results of the net change. [208]

8.6 WLI's ADAS-enhanced RUC data

In its 2004-05 modeling per the EPA and State MOU, the NDDH used ADAS-enhanced Rapid Update Cycle (RUC) data. The mesoscale RUC data were provided via contract by WindLogics, Inc. (WLI), St. Paul, Minnesota. WindLogics has described the advantages of the RUC model compared to forecast models, such as MM5. [209]

---

[204] See Exhibit 84, page 4.

[205] See section 5.2 in Addendum B to Exhibit 158.

[206] Id.

[207] The background concentration for sulfur dioxide will likely change in the future when better information or methods for estimating it become available.

[208] See Appendix B to Addendum B to Exhibit 158.

[209] See *RUC Analysis-based CALMET Meteorological Data for the State of North Dakota*, which is included in Addendum D to Exhibit 158, pages 1 through 5. "The RUC cycle is unique among the NCEP forecast systems in that analyses are produced every hour, versus every six hours for the other models used for longer term forecasting, ... The RUC cycle uses a process known as continuous assimilation, in which short, one-hour forecast segments are interspersed with applications of the data assimilation process. This means that each hour, the one-hour forecast fields are corrected, based on the real-time data collected by the National Oceanographic and Atmospheric Administration (NOAA). These

WindLogics has provided EPA Region 8 with the WindLogics-created meteorological output for years 2000, 2001 and 2002, and the NDDH understands that the FWS has a copy. [210]

On February 14, 2006, modeling practitioners with EPA Region 8, the NPS and the FWS orally indicated they would accept raw RUC data as input fields for the Calmet model. However, the coordinates of raw RUC data must be transformed into MM5-like coordinates as an initial step before use as input for Calmet. WindLogics used ADAS software (in Fortran) to perform the coordinate transformation, and its only modifications were added provisions to write output in a MM5-like format.

WindLogics provided a report that describes the software tools and the application of those tools to create the MM5-like input. [211] The ADAS software tools used by WindLogics were developed by the University of Oklahoma. These software tools were also used to complete assimilation of late-arriving NWS hourly surface observations during the RUC cycle for years 2000, 2001 and 2002. [212]

In its draft 2003 modeling, EPA Region 8 used 1994 MM5 data provided by contract with Alpine Geophysics [213] while also using 1992 MM5 data from an unknown source. In its draft 2003 modeling, the NDDH used 1990 MM4 and 1992 MM5 data initially provided by EPA, but subsequently provided by the federal FWS without MM5 model and MM5 model execution documentation. The 2002 MM5 data used for the model performance testing in this paper were also provided by the FWS – also without documentation.

Do the NOAA and University of Oklahoma ADAS models qualify as proprietary? No. EPA's modeling guideline at Appendix W to 40 CFR Part 51, section 3.1.1.c.vi states:

---

corrected fields serve as the starting point for the CALMET analysis described here." (Id., pages 2 and 3.) "Model first-guess fields derived from RUC archives have the advantage of being analyses rather than forecasts. That is they reflect the level of error in initial conditions used for the RUC model rather than the forecast errors." (Id., page 5)

[210] Personal telephone conversation between the NDDH and Mr. Tim Allen, FWS.

[211] See *RUC Analysis-based CALMET Meteorological Data for the State of North Dakota*, which is included in Addendum D to Exhibit 158, pages 5 through 12. One benefit of the ADAS software tools is the ARPS software method of interpolation of RUC data to a MM5-like grid. (Id., page 7.)

[212] See *A Comparison of NOAA RUC Analysis Surface Winds and ADAS-Enhanced RUC Analysis Winds with Surface Observations*, which also compares RUC to MM5 and is included in Addendum D to Exhibit 158. "A couple of limitations inherent in the RUC data files should be mentioned. First, the RUC system is a real-time system operating on a tight schedule. For this reason, late arriving observations may be missed in the RUC assimilation process. This is one of the key reasons for re-assimilating the observational data after the fact using ADAS. Second, the grid spacing available in the RUC archives is 40 km prior to 2003 ..." (Id., page 5.)

[213] See Exhibit 84, pages 4 – 7. See also Exhibit 85.

> "The developer must be willing to make the model available to users at reasonable cost or make it available for public access through the Internet or National Technical Information Service: the model cannot be proprietary."

This provision does not apply to RUC or ADAS because the language only refers to air quality dispersion models that are the subject of Appendix W. For example, MM5 is not listed and its use is not discussed in Appendix W.

Are WindLogics RUC data proprietary or confidential? No, [214] WindLogics license fee does not prevent third-party acquisition of the data. Since 2003, RUC data are available for purchase from another vendor, and for the NDDH's modeling domain the data purchase cost would be about 75% less. But then the vendor's RUC data must be transformed to MM5-like coordinates and formatted for use with Calmet, increasing costs.

8.7 Best choices in meteorological data and in its application

Meteorology is a critical component of modeling. Major sources in western North Dakota are widely separated and located more than 50 kilometers from PSD Class I areas. Time pathways of emitted sulfur dioxide between distant sources in the modeling domain and these areas can exceed 24 hours. Place (x, y and z) and time (t) variant meteorology across the modeling domain dominates model-estimated concentrations.

Weather conditions control transport and dilution of sulfur dioxide at and downwind of source discharge. [215] Modeled concentrations reflect these circumstances. Since pollution dilution is proportional to wind speed, the highest model-estimated and actual concentrations are due to low wind speeds along the plume transport path. [216] However, error in low wind speeds and corresponding wind reversals can cause significant error in modeled concentrations.

Better agreement between Calmet's layered output wind fields and wind observations results in better modeled plume transport and dispersion. Available observations for accuracy testing include wind measurements [217] and monitoring data. Existing studies indicate that the enhanced RUC2 data exhibit better agreement than MM5 data with wind observations. For example, WindLogics completed analyses of the accuracy of the meteorological output from

---

[214] See letter by Lawrence E. Volmert for Holland & Hart to Sara L. Laumann, EPA Region 8, dated June 5, 2003, at Tab B in Vol. 6 of Exhibit 95.

[215] For illustrations, see Appendix G of Addendum B and sections 4.0 and 10.2 of Addendum C to Exhibit 158.

[216] See results of accuracy tests of modeled concentrations using actual emissions and hourly CEM emissions in section 6 in Addendum C to Exhibit 158.

[217] These wind measurements were taken at wind-energy meteorological-tower locations in western North Dakota. The wind data have not been assimilated by the RUC or MM5 models.

application of those tools with NWS surface observations, [218] and ENSR completed accuracy analyses of NWS Bismarck upper air observations with RUC data and independent meteorological-tower wind data in western North Dakota with RUC data.

In addition, the NDDH has observed sulfur dioxide concentrations at several monitoring sites; better agreement (ratios) between estimated concentrations from Calpuff and observed concentrations from field monitors is desirable. When using current Calmet and Calpuff inputs, the ratios from accuracy tests are larger than 1.0, indicating that these models and all model data inputs are overestimating monitoring data and, thus, overestimating sulfur dioxide deterioration. [219]

8.8 EPA's major source SO_2 emission rates

Model-estimated sulfur dioxide concentrations are proportional to source-emitted sulfur dioxide; generally, on-the-ground concentrations double when a source's emission rate doubles. As illustrated in table 13, the sum total of major source sulfur dioxide emissions rates for EPA Region 8's methods of calculating these rates are substantially larger that the NDDH's rates. So model-estimated concentrations due to current sulfur dioxide emissions and estimated changes in concentrations due to changes in emissions after PSD baseline would be, other factors equivalent, substantially larger when using EPA Region 8's emissions data.

No "apples-to-apples" comparison of Calmet–Calpuff estimated concentrations using EPA Region 8's 90th percentile choices for emission rates and NDDH's "actual emission" rates is available. All other data inputs for this comparison would be identical so as to eliminate effects of differences in model inputs for mesoscale meteorology, model control files variables and model grid scales.

The NDDH and EPA Region 8 used their sulfur dioxide emissions inventories when modeling estimated concentrations or estimated changes in concentrations in North Dakota's PSD class I areas, respectively. Table 13 also illustrates the net change in sulfur dioxide emission rates after PSD baseline as the difference between current rates and baseline rates.

Region 8's estimates of sulfur dioxide PSD class I 24-hour increment consumption were developed by modeling increment-consuming and expanding emissions, which were calculated as the source-by-source difference between 2000-01 emissions and estimated 1976-77 emissions. And Region 8 used 1990-94 meteorology rather than meteorology concurrent with 2000-01 emissions. (Sulfur dioxide emissions for all source categories during years 1990-94 were larger than emissions during 2000-01. See figure 2.)

---

[218] See *A Comparison of NOAA RUC Analysis Surface Winds and ADAS-Enhanced RUC Analysis Winds with Surface Observations*, which also compares RUC to MM5 and is included in Addendum D to Exhibit 158.

[219] See, respectfully, sections 5 and 6 in Addendum C and Appendix B of Addendum B to Exhibit 158.

Table 13. Major Source Sulfur Dioxide Emission Rates		
Time lines	NDDH * (lb/op-hr)	EPA Region 8 ** (lb/hr)
Current time line (2000-01)	43,658	53,602
PSD baseline	29,684	26,352
Current minus baseline	13,974	27,250
Ratio: difference ÷ current	0.32	0.51
* Actual emission rates as defined by regulation. ** Ninetieth percentile emission rates as a deviate from its guidance.		

Table 14 provides data for actual sulfur dioxide concentrations and model-estimated increment consumption. Region 8's estimated sulfur dioxide 24-hour class I increment consumption for TRNP-SU ranges between 7.5 and 10.5 ug/m3, while the SH actual sulfur dioxide concentrations there for 2000 and 2001 were 9.4 and 8.8 ug/m3. So the SH actual sulfur dioxide concentrations are inside the range of Region 8's model-estimated increment consumption. Because Region 8's increment-affecting emissions were 51% of (not similar to) its 2000-01 emissions (see table 13), its model estimates of increment consumption are out of balance with actual sulfur dioxide concentrations.

8.9 Baseline concentration and allowable SO_2 emission rates

Implementing regulations define "baseline concentration" as "that ambient concentration level which exists in the baseline area at the time of the applicable minor source baseline date." A baseline concentration includes: "the actual emissions representative of sources in existence on the applicable minor source baseline date [except actual emissions from any major stationary source on which construction commenced after the major source baseline date]" and "the allowable emissions of major stationary sources which commenced construction before the major source baseline date but were not in operation by the applicable minor source baseline date." [220]

[220] See Addendum G to Exhibit 158, pages 12 and 13.

Table 14. Actual HSH 24-hour Sulfur Dioxide Concentrations and Model-estimated HSH 24-hour PSD Sulfur Dioxide Increment Consumption			
Modeled meteorology year	Actual HSH 24-hour concentrations (ug/m3)	Model-estimated HSH PSD 24-hour increment consumption (ug/m3)	
		NDDH MOU results report *	EPA Region 8 2003 draft report *
1990 / 91 / 92 / 93 / 94			
TRNP–SU	no data	1990 – 1994 meteorological data were not modeled	10.5 / 9.7 / 8.6 / 7.5 / 9.3
TRNP–NU	9.7 / 12.6 / 10.5 / 12.8 / 19.7		8.1 / 11.0 / 6.3 / 7.5 / 5.5
2000 / 2001 / 2002			
TRNP- SU	9.4 / 8.8 / 8.3	4.4 / 4.0 / 4.7 **	same as above (2000 – 2002 meteorological data were not modeled)
TRNP–NU	no data	3.6 / 3.6 / 4.5	
* Emissions were for years 2000 and 2001. ** Number for 2002 based on emissions for 2002-2003.			

When modeling, the baseline concentration for sulfur dioxide is determined from model-estimated concentrations using an inventory of baseline sulfur dioxide emissions. (See sections 4.1 and 6.2.) The sulfur dioxide emissions of sources in this inventory are defined by one of two options for "actual emissions" in implementing rules: (1) the average rate, in tons per year, at which the unit actually emitted the contaminant ... or (2) source-specific allowable emissions (assuming such are equivalent to the unit's actual emissions). [221]

Source-specific allowable emissions are incompatible with the implementing definition for baseline concentration, unless allowable sulfur dioxide emissions are equivalent to actually emitted sulfur dioxide. Power-plant sulfur dioxide allowable rates are a maximum 3-hour rolling average. The sulfur dioxide allowable rates during PSD baseline were substantially larger than actual peak baseline rates – 64,339 lb/hr and 35,165 lb/hr, respectively – using pre-2000 methods and source baseline coal data to calculate baseline emission rates for pre-2000 modeling. [222] And power-plant sulfur dioxide allowable emissions during the current period of

[221] See Exhibit 158, section 3.4.1.

[222] See Exhibit 133, section 2.1.1 and table 6.

2000-01 are also substantially larger than the sulfur dioxide actually emitted during that period as shown in table 15.

Table 15. Sulfur Dioxide Emission Rates for the Seven in-state Power Plants **			
Year	Current permit allowable * (lb/hr)	90th percentile of 24-hour block averages of hourly CEM data (lb/hr)	Annual average of hourly CEM data during operating hours (lb/op-hr)
2000	64,339	44,761.7	38,730.9
2001	64,339	45,509.6	38,921.6
* maximum 3-hour rolling average ** See Exhibit 33, table 3.			

8.10 Baseline concentration and tracking PSD increment consumption

In 1974, EPA established and adopted the "baseline concentration" during PSD baseline as the benchmark for tracking changes in air quality. "In 1978, EPA determined that it would track increment consumption: (1) by defining and modeling only increment consuming "allowable emissions" rather than a full emissions inventory of "allowable emissions" (or actual emissions); (2) by placing 'primary emphasis on tracking emission changes rather than establishing a baseline concentration,' and (3) by determining that it would 'no longer suggest that the baseline concentration be formally established.' " [223] (citation omitted)

In 1980, it changed the 1978 rule for "baseline concentration" and added a rule definition for "actual emissions". And, its 1978 augmented modeling method appeared in its guidance; i.e., it initiated modeling increment-affecting emissions rather than both PSD baseline and current inventories of emissions.

In 1990, however, EPA stated that the baseline concentration is "... the reference point for determining air quality deterioration in the area" without changing an emphasis on modeling increment-affecting emissions. [224] But in 2002, EPA reaffirmed its practice of using the paired-

[223] See Exhibit 158, page 16.

[224] See EPA's **draft** *1990 New Source Review Workshop Manual – Prevention of Significant Deterioration and Nonattainment Permitting*, page C.6.

in-space-and-time method of calculating deterioration without reference to the baseline concentration. [225]

Five options for tracking changes in sulfur dioxide concentrations after the PSD baseline date are summarized in table 16. Tracking changes requires a reference or benchmark concentration; increases over the benchmark are compared to a PSD increment. Or, the benchmark can be added to a PSD increment and the sum compared to current concentrations.

Paired-in-space-and-time method's practical flaw

Computer models are required to estimate deterioration or improvement in air quality since 1977, a date set by the federal CAA, since no reliable on-the-ground monitoring data for North Dakota are available for that time period. There are many methods to calculate likely air quality increases (deterioration) or decreases (improvement) using model estimates of actual concentrations due to PSD baseline emissions and to current emissions.

Various methods for analyses of observed and modeled sulfur dioxide concentrations are needed because the single values of CAA PSD increments do not adequately describe historic trends in concentrations. These analyses include model receptor network averages of model-estimated concentrations. Other examples are provided in section 10 of the MOU Protocol Results Report.

The CAA's text for PSD increments do not constrain use of those methods, nor does that text constrain deterioration of sulfur dioxide when FLMs certify no-adverse AQRV impacts. However, the text narrows the methods of calculating changes in air quality. For example, the State's MOU Protocol incorporates a baseline concentration as required by the CAA; [226] the method is dubbed the "paired-in-space[place]-only" method. The baseline concentration is the SH of short-term concentrations during the year at a receptor in a PSD class I area. So it is a worst-case concentration at the site of the receptor.

Another method of tracking air quality changes does not incorporate a "baseline concentration" and is dubbed the "paired-in-space-and-time" method. Since 1978 or 1980, modeling for PSD used increment-consuming and increment-expanding emissions, which are changes (increases or decreases) in emissions of existing sources after PSD baseline as well as emissions of new sources. The practice of modeling increment-consuming and increment-expanding emissions (EPA's 2002 and 2003 draft reports), rather than baseline and current emissions inventories, inhibits:

- Use of a baseline concentration for tracking changes in concentrations after PSD baseline. The baseline concentration can be estimated by modeling the baseline emissions inventory.

---

[225] See Exhibit 136, which is dated October 16, 2002 and is titled *Calculation of Prevention of Significant Deterioration (PSD) Increment Consumption and Expansion, Paired in Space and Time*.

[226] See also 45 FR (No. 154) page 52718, left column.

Table 16. Options for tracking changes in concentrations after the PSD baseline date.

	Reference for tracking changes in concentrations after the PSD baseline date			CAA frames increments in duration (e.g., short-term) and at location		
method for tracking changes in concentrations	PSD baseline benchmark [1]	comparison to rule-defined "baseline concentration"	current-period concentrations [2]	each short-term period (temporal) during the year [3]	each receptor (spatial) [3]	implementation
paired in space and time (PST)	for each receptor; each short-term period [4]	overstates occurrences of incr. consumption	for each receptor; each short-term period	yes	yes	NDDH – 1999 EPA R8 – 2002 & 2003
paired in space only (PSO)	for each receptor; the SH short-term period	matches, if each receptor represents a sub-scale portion of the class I area	for each receptor; each short-term period	yes	yes	NDDH – 2004/05
unpaired in space (HSH)	for the class I area; the HSH of all receptor SHs	understates occurrences of incr. consumption	for each receptor; each short-term period	yes	yes	no
unpaired in space (ASH)	for the class I area; the average of all receptor SHs	matches	for each receptor; each short-term period	yes	yes	no
unpaired in space (ASH)	for the class I area; the average of all receptor SHs	matches	for the class I area: the average of receptor concentrations for each period	yes	no	NDDH – 2002 & 2003 (as the MAAL)

[1] Determined by modeling an inventory of baseline emissions.
[2] Determined by modeling an inventory of current emissions.
[3] EPA recently stated: "[T]o determine compliance with the PSD increment, one should determine whether the net change in increment consuming emissions [sic] emissions since the baseline date has resulted [in changes] in pollutant concentrations exceeding the PSD increment at any specific time (temporal) and location (spatial) in the ***current*** year." (Exhibit 136, emphasis added.)
[4] Baseline concentrations were not explicitly determined, since only increment-affecting emission rates were modeled.

- Exploring deterioration or improvement in modeled worst-case concentrations.

- Testing the accuracy of model-estimated concentrations against actual concentrations, which requires modeling the current emissions inventory.

Paired-in-space-and-time method's technical flaw

When modeling baseline and current emission inventories, the annual meteorology data inputs for the modeling are the same for each inventory. EPA's current method of determining deterioration in time as well as in space (place) cannot be duplicated with monitoring data. Comparison of monitoring data on a Julian day during current period year(s) to monitoring data on the same Julian day during PSD baseline year(s) as a measure of deterioration presumes that the weather was the same on that day during both time periods. Identical weather on the same Julian day during both time periods is extremely unlikely.

EPA has preferred to model increment-affecting emissions to track deterioration, and model output from modeling increment-affecting emissions represents deterioration (or improvement) in concentrations. This technique implicitly forces time-pairing modeled current and PSD baseline sulfur dioxide concentrations at each Class-I-area receptor.

The State's MOU Protocol mimics EPA's methods by modeling current and PSD baseline inventories of emitted sulfur dioxide using the same meteorological data and then subtracting baseline concentrations from time-paired current concentrations at each receptor.

Deterioration that is calculated by the paired-in-space-and-time method is scientifically credible only when there are no changes in locations or pollutant-plume heights of sources following PSD baseline. However, calculated deterioration by this method is not credible if meteorological data were available and used for the baseline period or if there are changes in locations or plume heights of sources, because the models do not adequately duplicate real-world pollutant transport and dispersion between sources and receptors (see sections 5.12 and 6.4).

Deterioration that is calculated by using the baseline concentration for the Class I area, as defined by the CAA and EPA-implementing regulations and discussed in preambles, does not implicitly rely on an assumed robust quality of the models to estimate concentrations having acceptable day-to-day accuracy at sites of model receptors or monitors (see section 6.2).

8.11 Benefits of monitoring data in tracking PSD increment consumption

The NDDH has not used monitoring of actual concentrations in lieu of model-estimated concentration to estimate changes in concentrations since PSD baseline, including the State's periodic review process from 2000 through 2005. The NDDH has assessed annual trends in actual sulfur dioxide concentrations since 1977; however, monitoring technology advanced during the late 1970s and monitoring data during prior years were not reliable.

The NDDH also used actual concentrations of sulfur dioxide as indicators of Calmet–Calpuff performance through accuracy tests published in 1999, 2002, 2003 and 2004-05. Results of all tests have indicated that the State's protocols using these models overestimated actual concentrations. The NDDH improved the background concentration for sulfur dioxide and the mathematics of the test for the State's MOU Protocol.

Since 1980, several EPA documents have advocated model performance uncertainty analyses or accuracy testing. [227] Uncertainty or inaccuracy in model-estimated concentrations remains in spite of reasoned choices in model construction [228] (computer codes as approved by EPA) and model data inputs.

Modeled concentrations are – at best – estimates of actual concentrations. So modeled changes in air quality such as increases (deterioration) or decreases (improvement) in on-the-ground ambient sulfur dioxide using models are also estimates. Nevertheless, decision makers usually treat modeled estimates of changes as absolutes in administration of the CAA PSD class I increments, because methods for including error or bias correction of estimated changes in decision making have not been developed and adopted.

8.12 Model receptor grids for PSD class I areas

Subsequent to the NDDH's hearings, FLMs have recommended model receptor sets for the NU and SU of TRNP and for the LNWA wherein receptors are less than 1 km apart. No supporting documentation for reasons for these receptor sets has been published. For example, the apparent spatial variations within these class I areas for water and soil quality, wet and dry deposition, ozone, and flora and fauna do not justify a receptor spatial scaling less than 2 km. [229]

Adverse impact determinations do not rely solely on the highest or second-highest model-estimated concentrations among receptor sets, but instead consider the frequency, magnitude and duration of actual concentrations or model-estimated concentrations. [230] Distributions of

---

[227] See, as examples, these documents:

- *Interim Procedures for Evaluating Air Quality Models*, EPA's Office of Air Quality Planning and Standards, 1981.
- *Performance Measures and Standards for Air Quality Simulation Models*, EPA 450/4-79-032EPA, Office of Air Quality Planning and Standards, 1979.
- *Guideline on Air Quality Models*, Appendix W to 40 CFR Part 51, EPA's Office of Air Quality Planning and Standards, 1978 and subsequent updates.

[228] See Exhibit 122, *A User's Guide for the Calmet Meteorological Model*, and Exhibit 123, *A User's Guide for the Calpuff Dispersion Model*.

[229] See http:/biology.usgs.gov/npsveg/thro/metathroaa.html

[230] See Exhibit 18 titled *Federal Land Managers' Air Quality Related Values Workgroup (FLAG) Phase I Report*, page 7.

model-estimated concentrations at neighboring model receptors are likely not statistically independent. When this is so, useful accuracy in model-estimated concentrations is not improved by reducing the spacing of receptors in receptor sets or networks.

At the first of the NDDH's hearings in 2002, the NPS testified in support of a 2 km grid of receptors. [231] More recently (March 2006), the NPS agreed to the 2 km grid, as in the State's MOU Protocol, for AQRV assessments related to a proposed new power plant in western North Dakota. [232]

[231] See *Transcript of Hearing*, Exhibit 48, Vol. I, pages 144 and 145.

[232] See email from John Notar, NPS, to Terry L. O'Clair and others dated April 5, 2006, regarding *NPS comments on GNP [Great Northern Power] response*. "The National Park Service concurs with the facts and contents of the Great Northern Power Development L.P. March 9, 2006 letter."

Responses to Recurring Issues
Related to North Dakota's Computer Modeling
of Sulfur Dioxide in CAA PSD Class I Areas

ATTACHMENT A

Supplemental Data

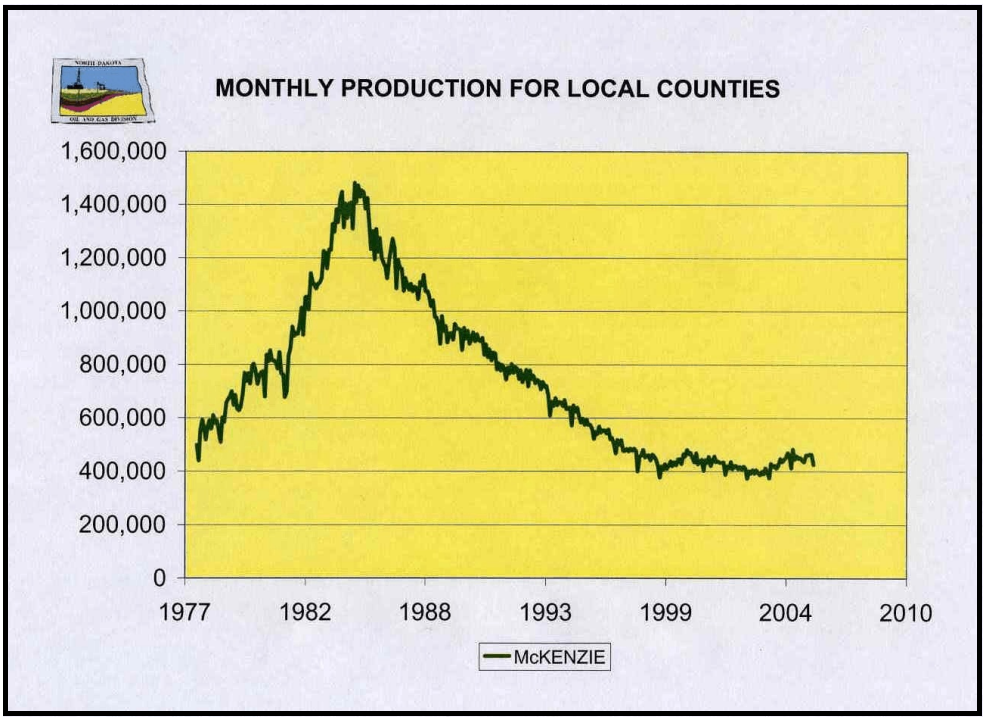


Figure A1. (See sections 2.4, 4.7 and 4.8.)

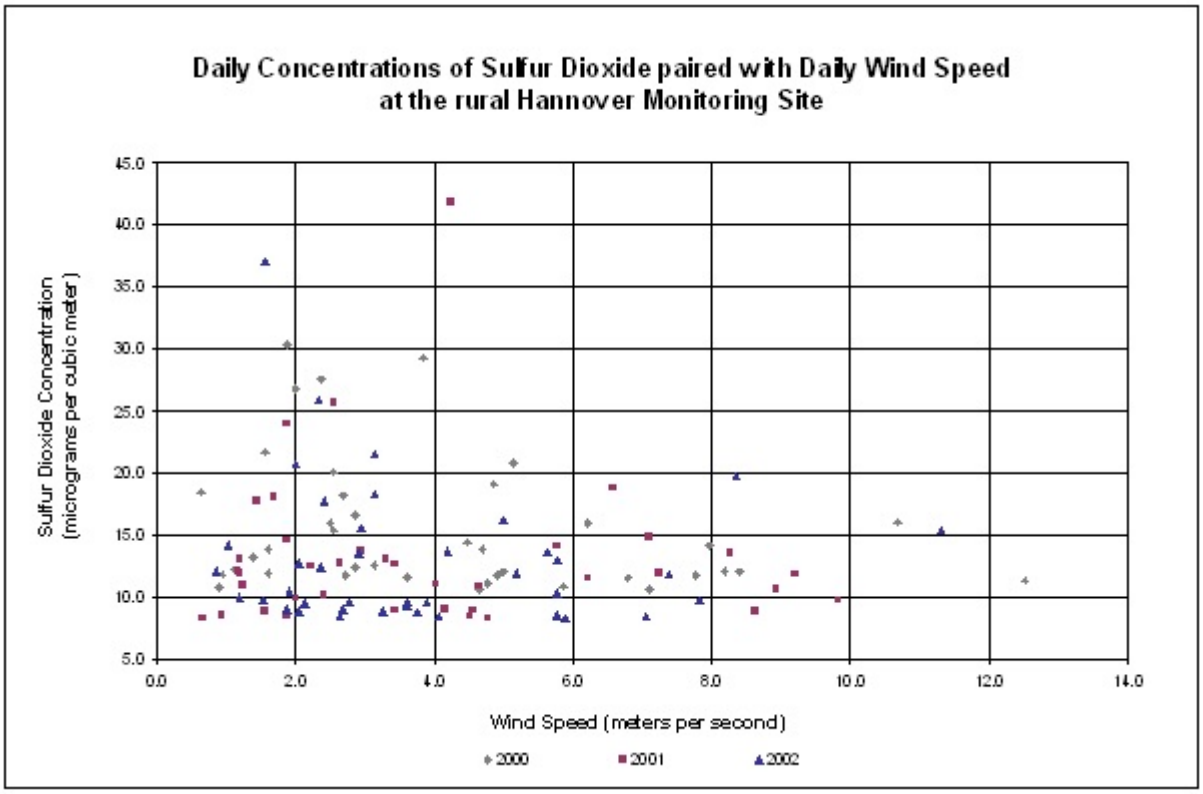


Figure A2. (See sections 2.8 and 5.12.)

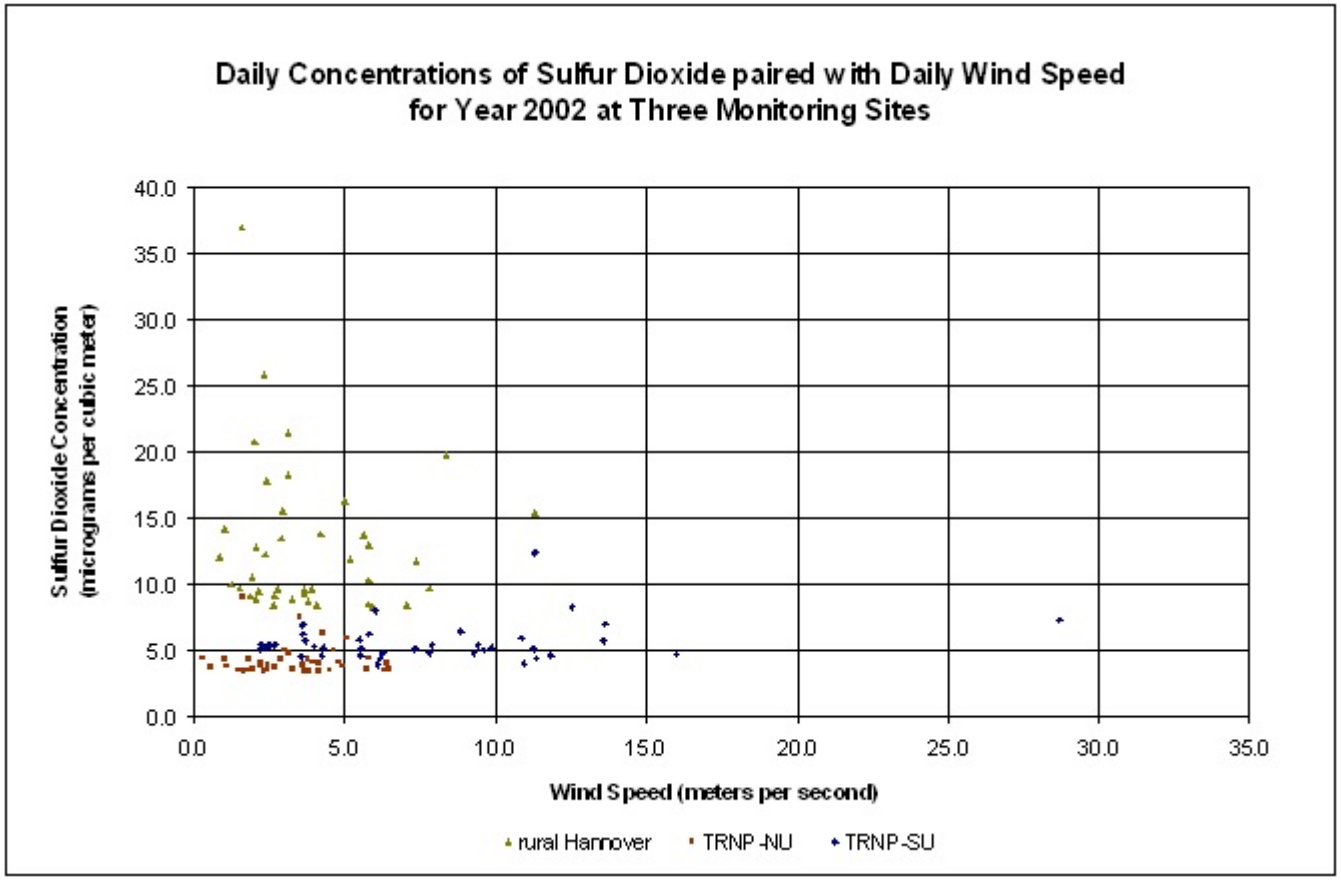


Figure A3. (See section 2.8 and 5.12.)

The average daily wind speeds at the TRNP-NU monitoring site for the 40 largest daily sulfur dioxide concentrations during year 2002 were lower than the average daily speeds at the TRNP-SU and the rural Hannover monitoring is sites.

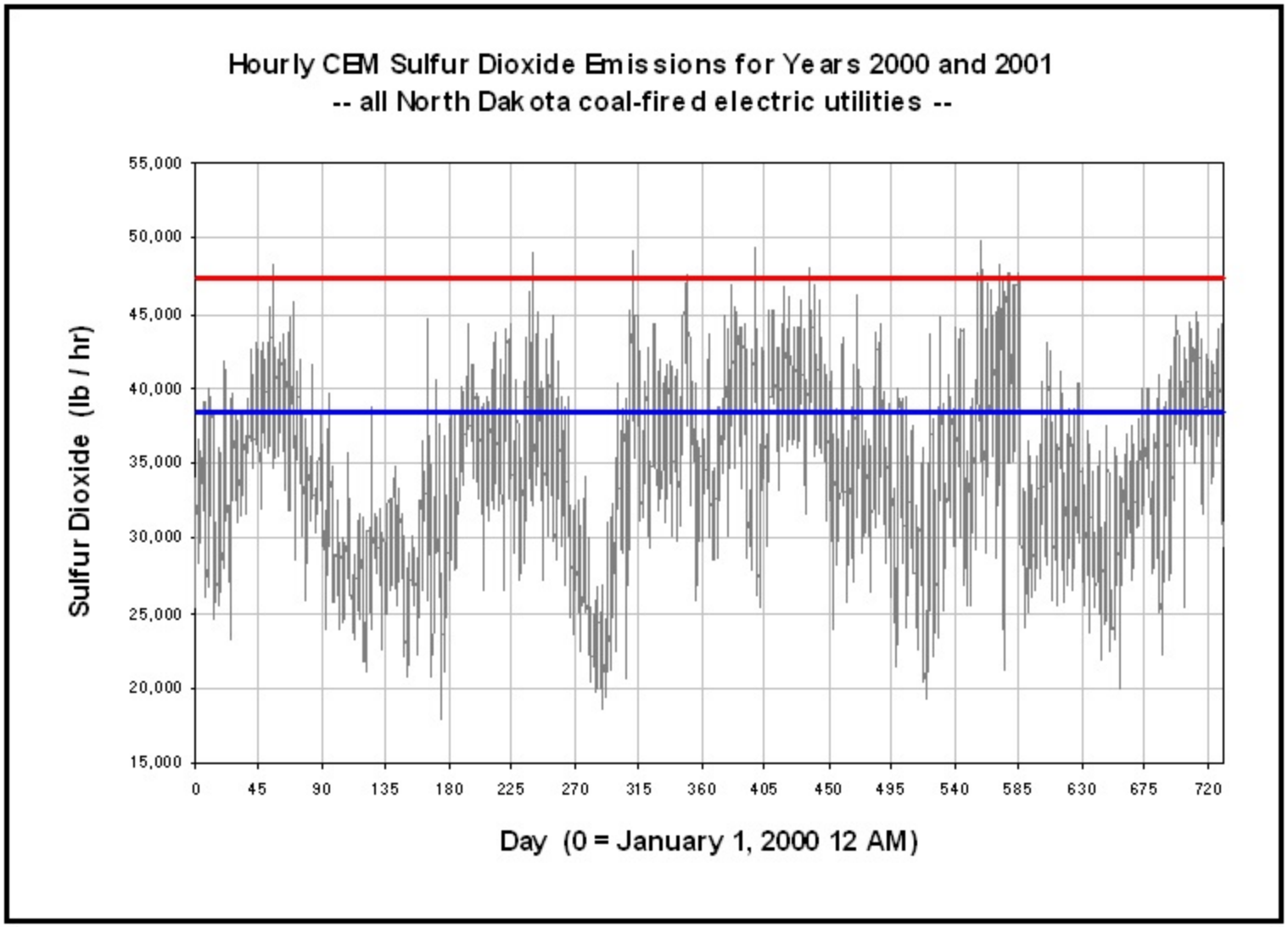


Figure A4. (The **red** bar is the sum of EPA Region 8's 90th percentiles of daily averaged hourly sulfur dioxide CEM emissions, and the **blue** bar is the sum of the NDDH's "actual emissions". See sections 4.2, 4.3 and 4.6.)

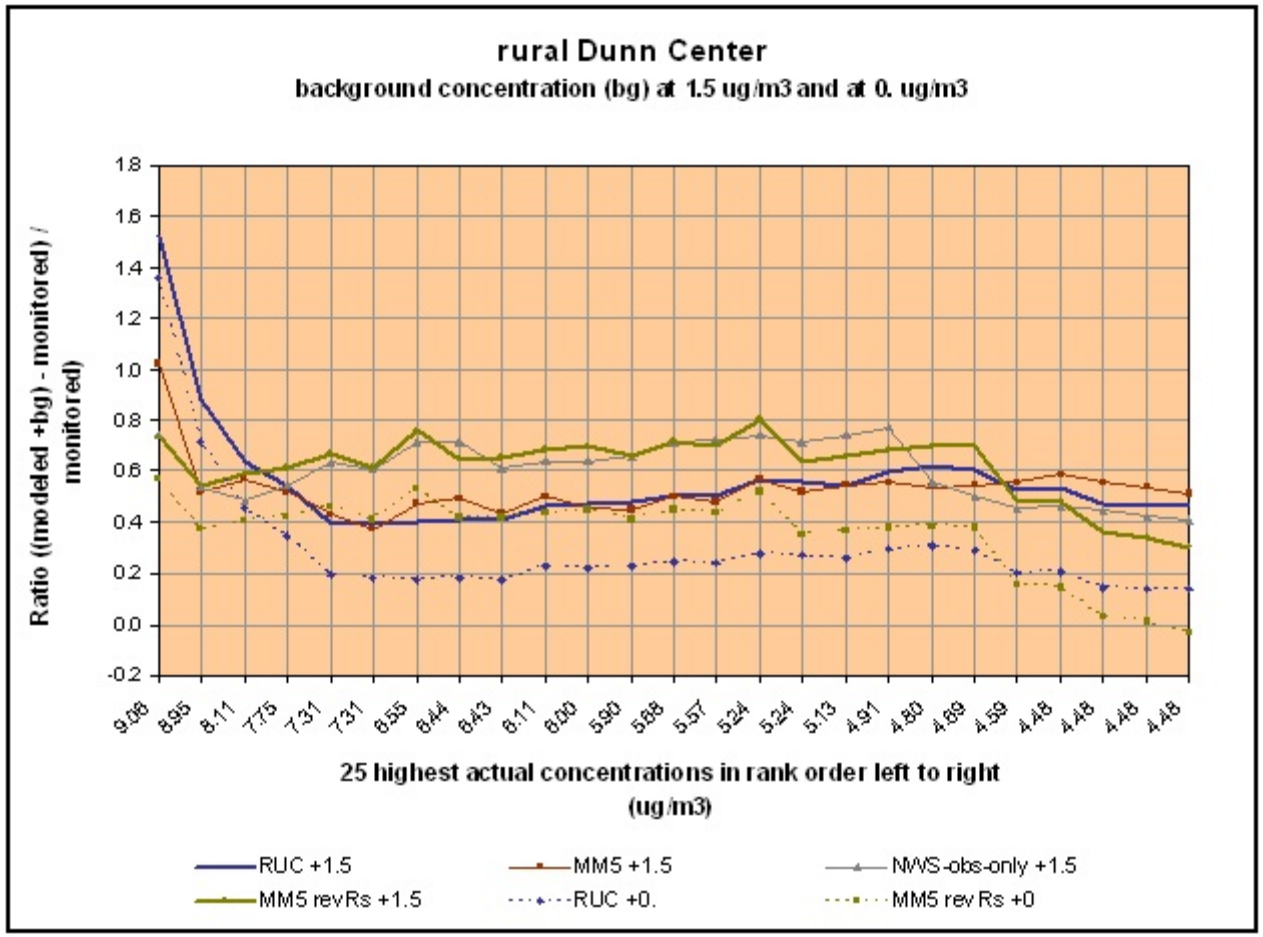


Figure A5. (See section 5.5.)

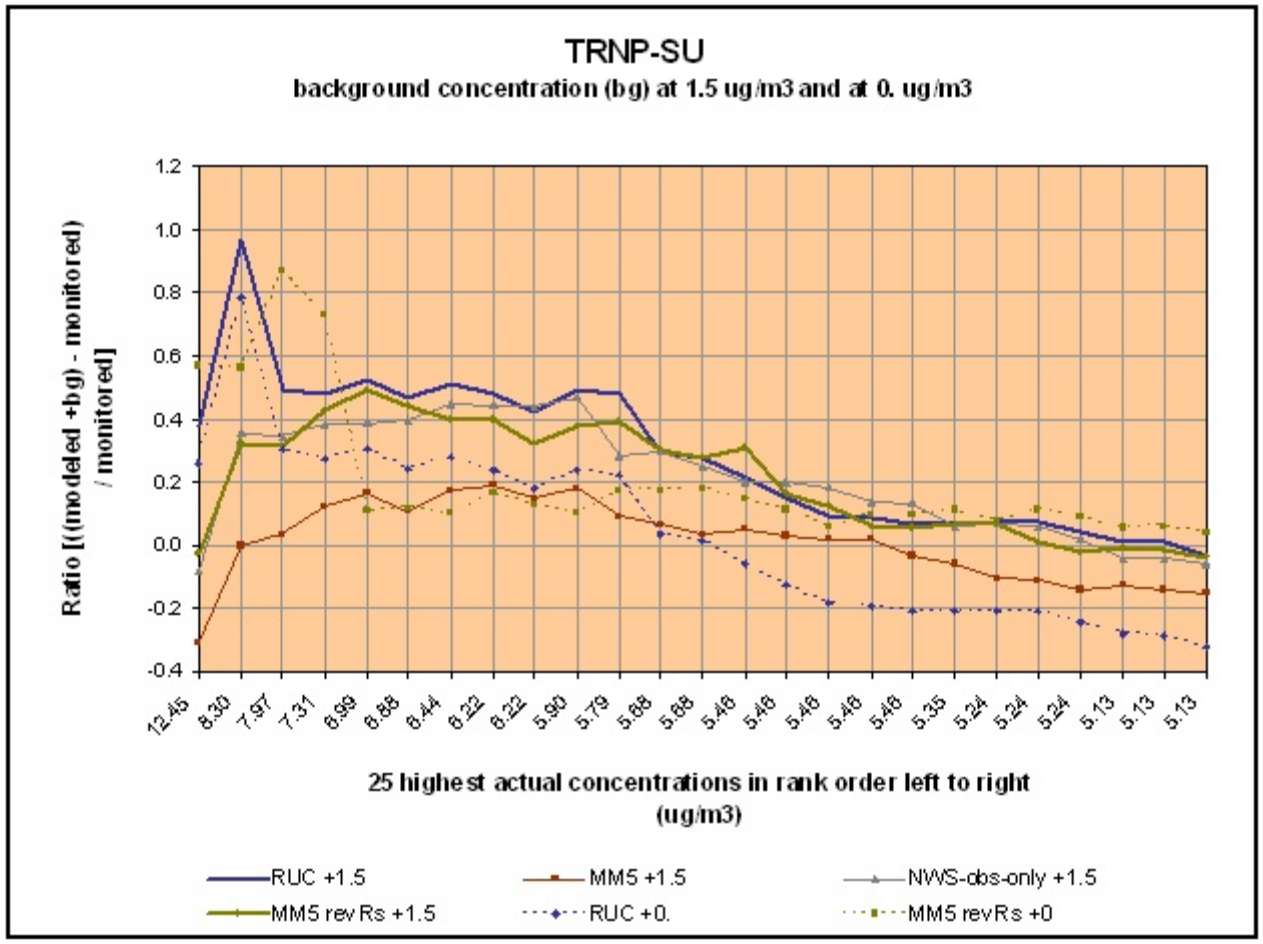


Figure A6. (See section 5.5.)

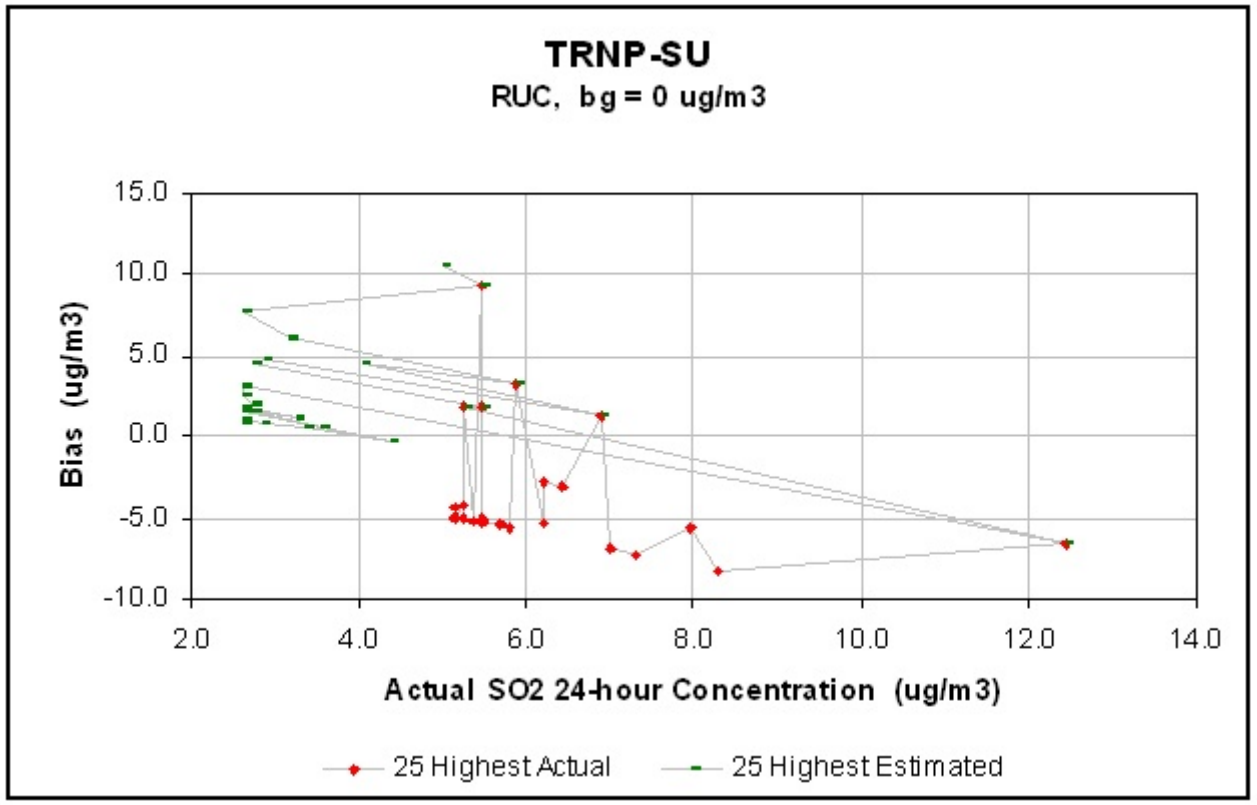


Figure A7. (See section 5.8.)

Responses to Recurring Issues
Related to North Dakota's Computer Modeling
of Sulfur Dioxide in CAA PSD Class I Areas

ATTACHMENT B

Model Accuracy Using

Hourly CEM SO_2 Emissions

Table B1. Error and Bias in Model-estimated * Sulfur Dioxide 24-hour Concentrations

	Largest 25 Actual Concentrations		Largest 25 Modeled Concentrations	
	RUC (1)	RUC (2)	RUC (1)	RUC (2)
rural Hannover				
ME (ug/m3)	9.6	9.9	7.8	7.6
MB (ug/m3)	-7.3	-8.0	6.6	6.6
TRNP-NU				
ME (ug/m3)	2.7	2.7	2.1	1.8
MB (ug/m3)	-1.4	-1.6	2.1	1.7
TRNP-SU				
ME (ug/m3)	5.0	4.9	3.2	2.9
MB (ug/m3)	-3.6	-3.6	2.7	2.9

* SO2 background concentration for model-estimated SO2 concentrations is 0 ug/m3.

Table B2. Correlation, Error and Bias in Model-estimated * Sulfur Dioxide 24-hour Concentrations

	RUC (1)		RUC (2)	
	N		N	
rural Hannover				
PCC		0.25		0.01
MNE	45	1.21	46	1.19
MNB		0.70		0.66
TRNP-NU				
PCC		0.21		0.24
MNE	41	0.84	40	0.76
MNB		0.50		0.42
TRNP-SU				
PCC		-0.23		-0.21
MNE	44	1.01	43	0.93
MNB		0.45		0.35

* SO2 background concentration for model-estimated SO2 concentrations is 1.5 ug/m3.

This attachment is an extension of model performance analysis and metrics in section 5.8. (See also sections 5.3, 5.5 and 5.7.) Here, sulfur dioxide hourly CEM emissions were paired with same-time Calmet output meteorological data, which dovetails the hour-to-hour variability of emissions and meteorology. Only RUC meteorological data were used in this performance analysis.

Column (1) in the tables at left: source sulfur dioxide emission rates are averages of annual tons (tons per year) during annual source operating hours. Data are from tables 8 and 9, respectively.

Column (2) in the tables at left: major source sulfur dioxide emission rates, except oil-field natural gas processing plants and the coal-to-synfuels plant, are hourly CEM emissions.

Graphs of normalized bias in model-estimated sulfur dioxide 24-hour concentrations time-paired with actual 24-hour concentrations are shown in figures B1, B2 and B3.

In sum, pairing hourly CEM sulfur dioxide emissions with hourly Calmet meteorology did not improve correlation between Calpuff-estimated sulfur dioxide 24-hour concentrations and actual 24-hour concentrations, nor did pairing reduce error and bias in Calpuff-estimated concentrations.

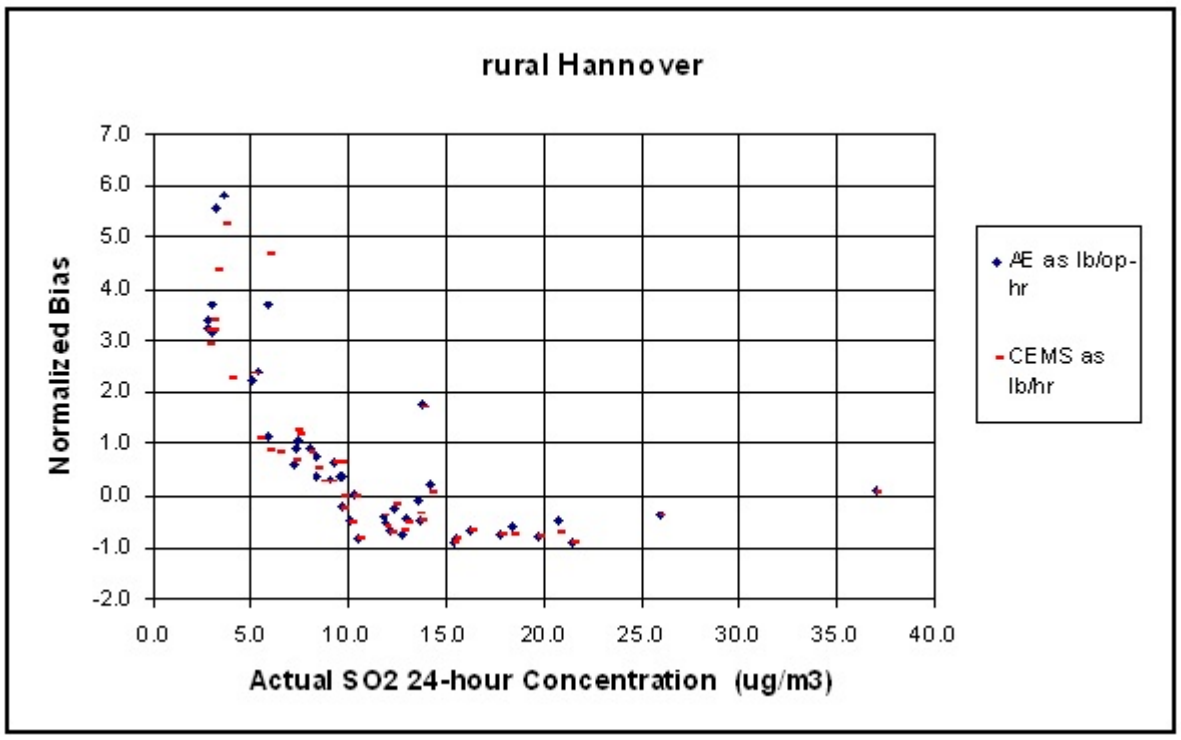


Figure B1.

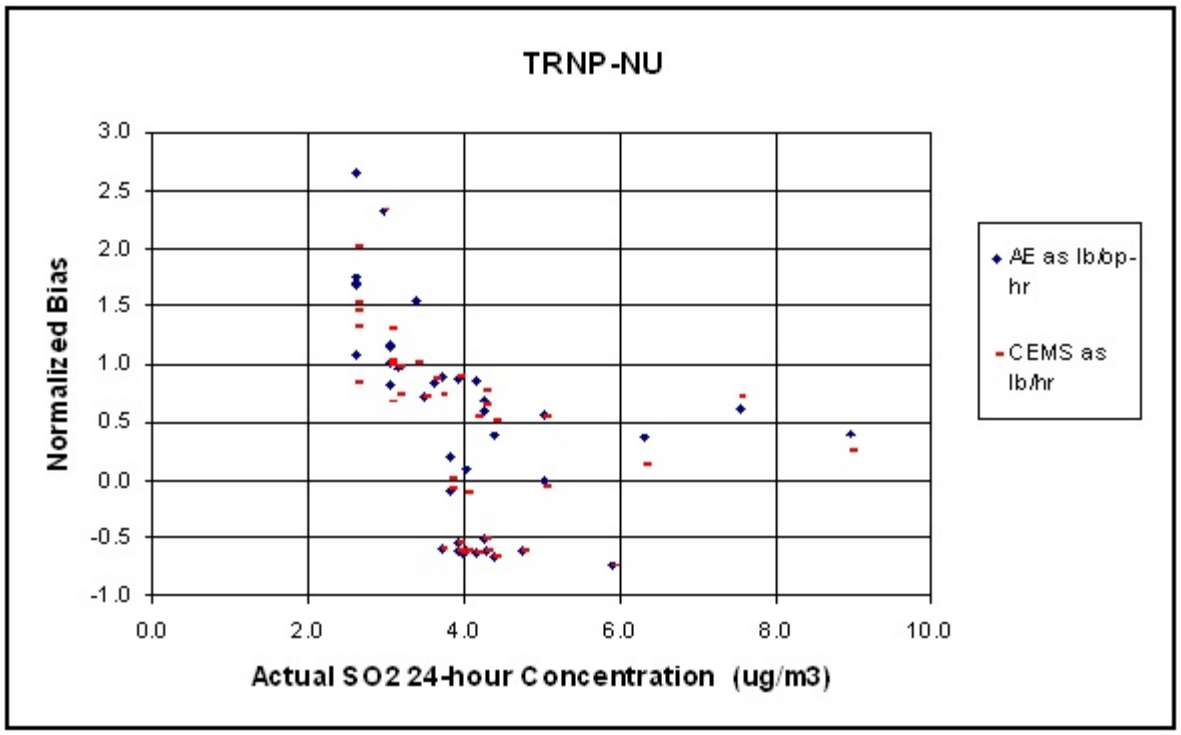


Figure B2.

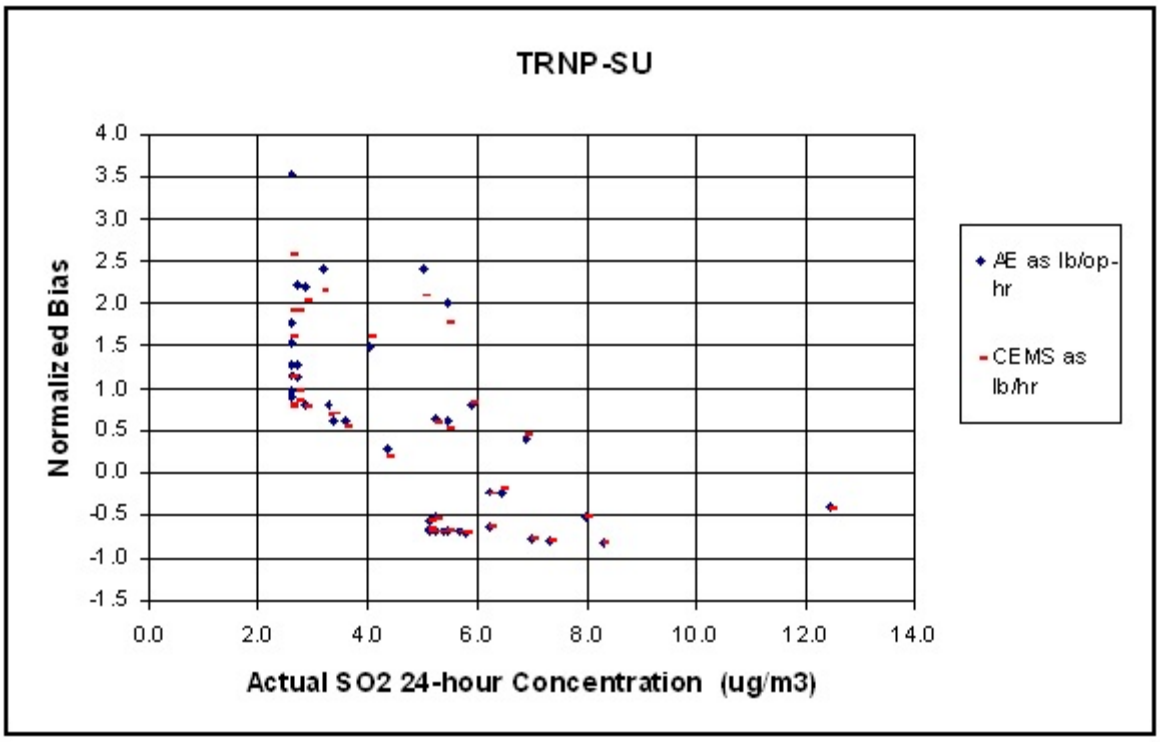

TRNP-SU
Normalized Bias
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
0.0
2.0
4.0
6.0
8.0
10.0
12.0
14.0
Actual SO2 24-hour Concentration (ug/m3)
AE as lb/op-hr
CEMS as lb/hr

Figure B3.

Responses to Recurring Issues
Related to North Dakota's Computer Modeling
of Sulfur Dioxide in CAA PSD Class I Areas

ATTACHMENT C

Source Contributions to

Model-estimated SO_2 Concentrations

The State's written MOU Protocol, as well as other NDDH and EPA Region 8 protocols since 1999 (section 3.10), use two categories of sources of sulfur dioxide emissions: the permitted industrial sources and the non-permitted oil and gas production flares and treaters. Oil and gas production sources having sulfur dioxide emission rates less than 0.001 gm/s [233] or located beyond 50 km from the TRNP PSD class I areas were not included. [234] Nearly all currently-operating permitted industrial sources are located at distances greater than 50 km from PSD class I areas (figure 1, page 3).

The State's written MOU Protocol also provided the sulfur dioxide emission rate for each industrial source. However, the protocol only provided the sum of rates for oil and gas production sources within 50 km of TRNP class I areas. [235] The sum of current (2000-01) sulfur dioxide emission rates by permitted industrial sources is 43,658.2 lb/op-hr (table 2, pages 34 and 35). The sum of current (2000) rates by oil and gas production sources is 1,487.4 lb/hr (table 4, page 38).

Calmet modeling using RUC meteorological data and Calpuff modeling using current sulfur dioxide emissions was completed as described in section 5.3. The contributions of oil and gas production flares and treaters to the 40 largest model-estimated sulfur dioxide 24-hour concentrations at model receptors co-located at sites of ambient monitors in the SU and NU of TRNP (figures 28 and 29, pages 84 and 85) and at rural Dunn Center (figure 5, page 7) are shown in figure C1. The oil and gas production sources did not contribute in amounts larger than 0.005 ug/m3 to model-estimated concentrations at the rural Hannover monitoring site.

In sum, contributions by oil and gas productions sources exceeded 10% of the model-estimated sulfur dioxide 24-hour concentrations: 4 of the 40 largest concentrations at the SU of TRNP monitoring site, 20 of the 40 largest concentrations at the NU monitoring site and 6 of the 40 largest concentrations at rural Dunn Center monitoring site.

The rural Dunn Center site lies between the TRNP and all permitted industrial sources but one; so contributions to a model-estimated concentration by the permitted industrial sources and by oil and gas production sources during a 24-hour period indicates changing wind directions during the 24-hour period or a wind direction reversal.

---

[233] 1 gm/s (grams/second) = 7.936 lb/hr

[234] A map of the permitted industrial sources is shown on page 18 of the MOU Protocol (Addendum B to Exhibit 159), and maps of oil and gas production sources are shown on page 54. A map of oil and gas production sources in the inventory and within 50 km of the LNWA was not included in the protocol.

[235] A table of sulfur dioxide emission rates for permitted industrial sources is shown on page 19 of the MOU protocol, and a table of the sum of oil and gas production sources' emission rates is shown on page 53.

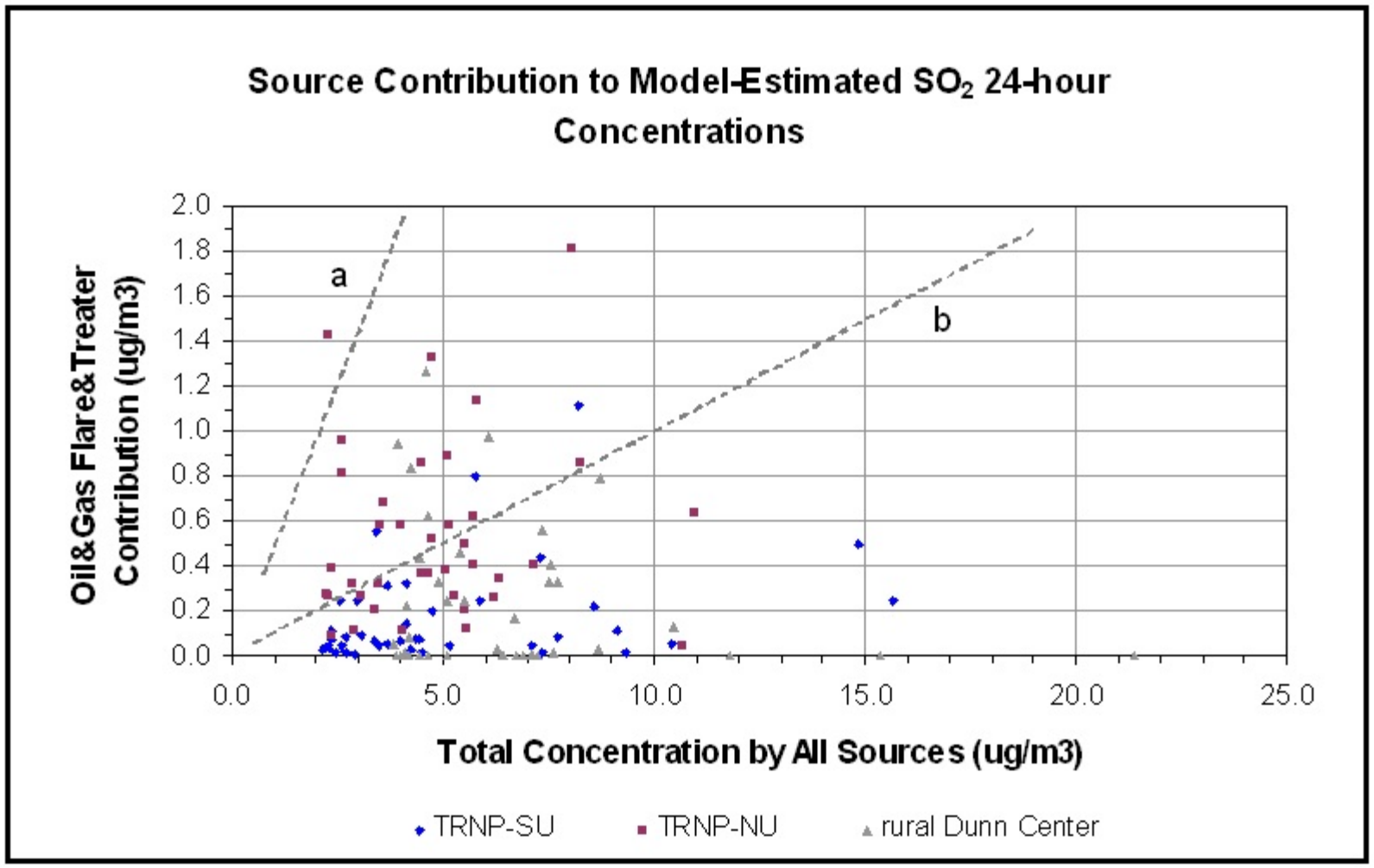


Figure C1. Line "b" represents oil and gas contributions that are 10% of total model-estimated SO_2 24-hour concentrations, and line "a" represents contributions that are 50% of total concentrations. For example, data points below the "b" line indicate that oil and gas production sources contributed less than 10% of total concentrations.

B L A N K P A G E